

SEQUENCE LISTING

<110> Hitoshi, Yasumichi
Jenkins, Yonchu
Rigel Pharmaceuticals, Inc.

<120> Methods of Assaying for Cell Cycle Modulators

<130> 021044-003310PC

<140> US 10/510,903

<141> 2004-10-08

<150> US 10/123,568

<151> 2002-04-15

<150> US 10/123,731

<151> 2002-04-15

<150> US 60/373,366

<151> 2002-04-16

<150> WO PCT/US03/11867

<151> 2003-04-15

<160> 60

<170> PatentIn Ver. 2.1

<210> 1

<211> 3599

<212> DNA

<213> Homo sapiens

<220>

<223> BRCA-1-Associated Protein-1 (BAP-1)

<400> 1

```

gcccgttgtc tgtgtgtggg actgaggggc cccggggggcg gtggggggctc ccggtggggg 60
cagcgggtggg gagggagggc ctggacatgg cgctgagggg ccgccccgcg ggaagatgaa 120
taagggtctgg ctggagctgg agagcgaccc aggcctcttc accctgctcg tggagattt 180
cgggtgtcaag ggggtgcaag tggaggagat ctacgacctt cagagcaaatt gtcagggccc 240
tgtatatgga tttatcttcc tgttcaaatt gatcgaagag cgccgggtccc ggcgaaagggt 300
ctctaccttg gtggatgata cgtccgtgat tgatgatgat attgtgaata acatgttctt 360
tgcccaccag ctgataccca actcttgtgc aactcatgcc ttgctgagcg tgctcctgaa 420
ctgcagcagc gtggacctgg gacccaccct gagtcgcatg aaggacttca ccaagggttt 480
cagccctgag agcaaaggat atgcgattgg caatgccccg gagttggcca aggcccataa 540
tagccatgcc aggcccgagc cagccaccct ccctgagaag cagaatggcc ttagtgagc 600
gcgaccatgg gaggcgttcc actttgtcag ctatgtgcct atcacaggcc ggctctttga 660
gctggatggg ctgaagggtct accccattga ccatggggccc tggggggagg acgaggagt 720
gacagacaag gcccggcggg tcatcatgga gcgtatcgcc ctgcgacctg caggggagcc 780
ctaccacgac atccgcttca acctgatggc agtggtgccc gaccgcagga tcaagtatga 840
ggccaggctg catgtgctga aggtgaaccg tcagacagta ctagaggctc tgcagcagct 900
gataagagta acacagccag agctgattca gaccacaaag tctcaagagt cacagctgcc 960
tgaggagtc aagtcagcca gcaacaagtc cccgctggtg ctggaagcaa acagggcccc 1020
tgcagcctct gagggcaacc acacagatgg tgcagaggag gcggctggtt catgcgca 1080
agccccatcc cacagccctc ccaacaaacc caagctagtg gtgaagcctc caggcagcag 1140
cctcaatggg gttcacccca accccactcc cattgtccag cggctgccgg cctttctaga 1200
caatcacaat tatgccaaat ccccatgca ggaggaagaa gacctggcgg caggtgtggg 1260
ccgcagccga gttccagtc gccacccca gcagtactca gatgatgagg atgactatga 1320
ggatgacgag gaggatgacg tgcagaacac caactctgcc cttaggtata aggggaaggg 1380

```

```

aacaggggaag ccagggggcat tgagcgggttc tgctgatggg caactgtcag tgctgcagcc 1440
caacaccatc aacgtcttgg ctgagaagct caaagagtcc cagaaggacc tctcaattcc 1500
tctgtccatc aagactagca gcggggctgg gagtccggct gtggcagtgc ccacacactc 1560
gcagccctca cccaccccca gcaatgagag tacagacacg gcctctgaga tcggcagtgc 1620
tttcaactcg ccaactgcgt cgctatccg ctcagccaac ccgacgcggc cctccagccc 1680
tgtcacctcc cacatctcca aggtgctttt tggagaggat gacagcctgc tgcgtgttga 1740
ctgcatacgc tacaaccgtg ctgtccgtga tctgggtcct gtcacagca caggcctgct 1800
gcacctggct gaggatggg tgctgagtc cctggcgctg acagagggtg ggaagggttc 1860
ctcgccctcc atcagaccaa tccaaggcag ccaggggtcc agcagcccag tggagaagga 1920
ggctcgtgaa gccacggaca gcagagagaa gacggggatg gtgaggcctg gcgagccctt 1980
gagtggggag aaatactcac ccaaggagct gctggcactg ctgaagtgtg tggaggctga 2040
gattgcaaac tatgaggcgt gcctcaagga ggaggtagag aagaggaaga agttcaagat 2100
tgatgaccag agaaggaccc acaactacga tgagttcatc tgcaccttta tctccatgct 2160
ggctcaggaa ggcatgctgg ccaacctagt ggagcagaac atctccgtgc ggcggcgcca 2220
aggggtcagc atcgccggc tccacaagca gcggaagcct gaccggcgga aacgctctcg 2280
cccctacaag gccaaagcgc agtgaggact gctggccctg actctgcagc ccactcttgc 2340
cgtgtggccc tcaccagggt ccttccctgc cccacttccc cttttcccag tattactgaa 2400
tagtcccagc tggagagtcc aggccctggg aatgggagga accaggccac attccttcca 2460
tcgtgccctg aggcctgaca cggcagatca gccccatagt gctcaggagg cagcatctgg 2520
agttggggca cagcgaggta ctgcagcttc ctccacagcc ggctgtggag cagcaggacc 2580
tggcccttct gcctgggcag cagaatatat attttaccta tcagagacat ctatttttct 2640
gggctccaac ccaacatgcc accatgttga cataagttcc tacctgacta tgctttctct 2700
cctaggagct gtctcgttgg gccaggtcc ttgtatcatg ccacggtccc aactacaggg 2760
tcctagctgg gggcctgggt gggccctggg ctctgggccc tgctgctcta gcccagcca 2820
ccagcctgtc cctgttgtaa ggaagccagg tcttctctct tcattcctct taggagagt 2880
ccaaactcag ggaccagca ctgggctggg ttgggagtag ggtgtcccag tggggttggg 2940
gtgagcaggc tgctgggatc ccatggcctg agcagagcat gtgggaactg ttcagtggcc 3000
tgtgaactgt cttccttgtt ctagccaggc tgttcaagac tgctctccat agcaaggttc 3060
tagggctctt cgccttcagt gttgtggccc tagctatggg cctaaattgg gctctaggtc 3120
tctgtccctg gcgcttgagg ctcaagaag cctctgtcca gcccctcagt attaccatgt 3180
ctccctctca ggggtagcag agacagggtt gcttatagga agctggcacc actcagctct 3240
tctgtctact ccagtttcct cagcctctgc aaggcactca ggggtggggga cagcaggatc 3300
aagacaaccc gttggagccc ctgtgttcca gaggacctga tgccaagggg taatgggccc 3360
agcagtgcct ctggagccca ggccccaaca cagccccatg gcctctgcca gatggctttg 3420
aaaaagggtg tccaagcagg cccctttatc tgtacatagt gactgagtgg ggggtgctgg 3480
caagtgtggc agctgcctct gggctgagca cagcttgacc cctctagccc ctgtaaatac 3540
tggatcaatg aatgaataaa actctcctaa gaatctcctg agaaaaaaaa aaaaaaaaaa 3599

```

<210> 2

<211> 729

<212> PRT

<213> Homo sapiens

<220>

<223> BRCA-1-Associated Protein-1 (BAP-1)

<400> 2

Met Asn Lys Gly Trp Leu Glu Leu Glu Ser Asp Pro Gly Leu Phe Thr
1 5 10 15

Leu Leu Val Glu Asp Phe Gly Val Lys Gly Val Gln Val Glu Glu Ile
20 25 30

Tyr Asp Leu Gln Ser Lys Cys Gln Gly Pro Val Tyr Gly Phe Ile Phe
35 40 45

Leu Phe Lys Trp Ile Glu Glu Arg Arg Ser Arg Arg Lys Val Ser Thr
50 55 60

Leu	Val	Asp	Asp	Thr	Ser	Val	Ile	Asp	Asp	Asp	Ile	Val	Asn	Asn	Met	65	70	75	80
Phe	Phe	Ala	His	Gln	Leu	Ile	Pro	Asn	Ser	Cys	Ala	Thr	His	Ala	Leu	85	90	95	
Leu	Ser	Val	Leu	Leu	Asn	Cys	Ser	Ser	Val	Asp	Leu	Gly	Pro	Thr	Leu	100	105	110	
Ser	Arg	Met	Lys	Asp	Phe	Thr	Lys	Gly	Phe	Ser	Pro	Glu	Ser	Lys	Gly	115	120	125	
Tyr	Ala	Ile	Gly	Asn	Ala	Pro	Glu	Leu	Ala	Lys	Ala	His	Asn	Ser	His	130	135	140	
Ala	Arg	Pro	Glu	Pro	Arg	His	Leu	Pro	Glu	Lys	Gln	Asn	Gly	Leu	Ser	145	150	155	160
Ala	Val	Arg	Thr	Met	Glu	Ala	Phe	His	Phe	Val	Ser	Tyr	Val	Pro	Ile	165	170	175	
Thr	Gly	Arg	Leu	Phe	Glu	Leu	Asp	Gly	Leu	Lys	Val	Tyr	Pro	Ile	Asp	180	185	190	
His	Gly	Pro	Trp	Gly	Glu	Asp	Glu	Glu	Trp	Thr	Asp	Lys	Ala	Arg	Arg	195	200	205	
Val	Ile	Met	Glu	Arg	Ile	Gly	Leu	Ala	Thr	Ala	Gly	Glu	Pro	Tyr	His	210	215	220	
Asp	Ile	Arg	Phe	Asn	Leu	Met	Ala	Val	Val	Pro	Asp	Arg	Arg	Ile	Lys	225	230	235	240
Tyr	Glu	Ala	Arg	Leu	His	Val	Leu	Lys	Val	Asn	Arg	Gln	Thr	Val	Leu	245	250	255	
Glu	Ala	Leu	Gln	Gln	Leu	Ile	Arg	Val	Thr	Gln	Pro	Glu	Leu	Ile	Gln	260	265	270	
Thr	His	Lys	Ser	Gln	Glu	Ser	Gln	Leu	Pro	Glu	Glu	Ser	Lys	Ser	Ala	275	280	285	
Ser	Asn	Lys	Ser	Pro	Leu	Val	Leu	Glu	Ala	Asn	Arg	Ala	Pro	Ala	Ala	290	295	300	
Ser	Glu	Gly	Asn	His	Thr	Asp	Gly	Ala	Glu	Glu	Ala	Ala	Gly	Ser	Cys	305	310	315	320
Ala	Gln	Ala	Pro	Ser	His	Ser	Pro	Pro	Asn	Lys	Pro	Lys	Leu	Val	Val	325	330	335	
Lys	Pro	Pro	Gly	Ser	Ser	Leu	Asn	Gly	Val	His	Pro	Asn	Pro	Thr	Pro	340	345	350	
Ile	Val	Gln	Arg	Leu	Pro	Ala	Phe	Leu	Asp	Asn	His	Asn	Tyr	Ala	Lys	355	360	365	
Ser	Pro	Met	Gln	Glu	Glu	Glu	Asp	Leu	Ala	Ala	Gly	Val	Gly	Arg	Ser	370	375	380	

Arg	Val	Pro	Val	Arg	Pro	Pro	Gln	Gln	Tyr	Ser	Asp	Asp	Glu	Asp	Asp	385	390	395	400
Tyr	Glu	Asp	Asp	Glu	Glu	Asp	Asp	Val	Gln	Asn	Thr	Asn	Ser	Ala	Leu	405	410	415	
Arg	Tyr	Lys	Gly	Lys	Gly	Thr	Gly	Lys	Pro	Gly	Ala	Leu	Ser	Gly	Ser	420	425	430	
Ala	Asp	Gly	Gln	Leu	Ser	Val	Leu	Gln	Pro	Asn	Thr	Ile	Asn	Val	Leu	435	440	445	
Ala	Glu	Lys	Leu	Lys	Glu	Ser	Gln	Lys	Asp	Leu	Ser	Ile	Pro	Leu	Ser	450	455	460	
Ile	Lys	Thr	Ser	Ser	Gly	Ala	Gly	Ser	Pro	Ala	Val	Ala	Val	Pro	Thr	465	470	475	480
His	Ser	Gln	Pro	Ser	Pro	Thr	Pro	Ser	Asn	Glu	Ser	Thr	Asp	Thr	Ala	485	490	495	
Ser	Glu	Ile	Gly	Ser	Ala	Phe	Asn	Ser	Pro	Leu	Arg	Ser	Pro	Ile	Arg	500	505	510	
Ser	Ala	Asn	Pro	Thr	Arg	Pro	Ser	Ser	Pro	Val	Thr	Ser	His	Ile	Ser	515	520	525	
Lys	Val	Leu	Phe	Gly	Glu	Asp	Asp	Ser	Leu	Leu	Arg	Val	Asp	Cys	Ile	530	535	540	
Arg	Tyr	Asn	Arg	Ala	Val	Arg	Asp	Leu	Gly	Pro	Val	Ile	Ser	Thr	Gly	545	550	555	560
Leu	Leu	His	Leu	Ala	Glu	Asp	Gly	Val	Leu	Ser	Pro	Leu	Ala	Leu	Thr	565	570	575	
Glu	Gly	Gly	Lys	Gly	Ser	Ser	Pro	Ser	Ile	Arg	Pro	Ile	Gln	Gly	Ser	580	585	590	
Gln	Gly	Ser	Ser	Ser	Pro	Val	Glu	Lys	Glu	Val	Val	Glu	Ala	Thr	Asp	595	600	605	
Ser	Arg	Glu	Lys	Thr	Gly	Met	Val	Arg	Pro	Gly	Glu	Pro	Leu	Ser	Gly	610	615	620	
Glu	Lys	Tyr	Ser	Pro	Lys	Glu	Leu	Leu	Ala	Leu	Leu	Lys	Cys	Val	Glu	625	630	635	640
Ala	Glu	Ile	Ala	Asn	Tyr	Glu	Ala	Cys	Leu	Lys	Glu	Glu	Val	Glu	Lys	645	650	655	
Arg	Lys	Lys	Phe	Lys	Ile	Asp	Asp	Gln	Arg	Arg	Thr	His	Asn	Tyr	Asp	660	665	670	
Glu	Phe	Ile	Cys	Thr	Phe	Ile	Ser	Met	Leu	Ala	Gln	Glu	Gly	Met	Leu	675	680	685	
Ala	Asn	Leu	Val	Glu	Gln	Asn	Ile	Ser	Val	Arg	Arg	Arg	Gln	Gly	Val	690	695	700	

Ser Ile Gly Arg Leu His Lys Gln Arg Lys Pro Asp Arg Arg Lys Arg
705 710 715 720

Ser Arg Pro Tyr Lys Ala Lys Arg Gln
725

<210> 3
<211> 3828
<212> DNA
<213> Homo sapiens

<220>
<223> nuclear zinc finger protein 95 (NP95)

<400> 3
cgactcctta gagcatggca tggctcagag gtgctggtaa aactgatggg ggtttttggct 60
gtccctcccc tcagcgccga caccatgtgg atccagggtc ggaccatgga cgggaggcag 120
acccacacgg tggactcgct gtccaggctg accaagggtg aggagctgag gcggaagatc 180
caggagctgt tccacgtgga gccaggcctg cagaggtgtg tctacagggg caaacagatg 240
gaggcggcc atacctctt cgactacgag gtccgcctga atgacaccat ccagctcctg 300
gtccgcccaga gctcgtgct cccccacagc accaaggagc gggactccga gctctccgac 360
accgactccg gctgctgcct gggccagagt gagtcagaca agtcctccac ccacggcgag 420
gcgccgccc agactgacag caggccagcc gatgaggaca tgtgggatga gacggaattg 480
gggctgtaca aggtcaatga gtacgtcgat gctcgggaca cgaacatggg ggcgtgggtt 540
gaggcgcagg tggtcagggt gacgcggaag gccccctccc gggacgagcc ctgcagctcc 600
acgtccaggc cggcgtgga ggaggacgtc atttaccacg tgaaatacga cgactaccgc 660
gagaacggcg tggtcagat gaactccagg gacgtccgag cgcgcgcccc caccatcatc 720
aagtggcagg acctggagggt gggccagggtg gtcattgtca actacaaccc cgacaacccc 780
aaggagcggg gcttctggtg cgacgcggag atctccagga agcgcgagac caggacggcg 840
cgggaactct acgccaacgt ggtgctgggg gatgattctc tgaacgactg tcggatcatc 900
ttcgtggacg aagtcttcaa gattgagcg cgggtgaag ggagcccat gggtgacaac 960
cccatgagac ggaagagcgg gccgtcctgc aagcactgca aggacgacgt gaacagactc 1020
tgccgggtct gcgcctgcca cctgtgcggg ggccggcagg accccgacaa gcagctcatg 1080
tgcatgagt gcgacatggc cttccacatc tactgcctgg acccgcccct cagcagtgtt 1140
cccagcgagg acgagtggta ctgccctgag tgccggaatg atgccagcga ggtggtactg 1200
gcgggagagc ggctgagaga gagcaagaag aaggcgaaga tggcctcggc cacatcgctc 1260
tcacagcggg actggggcaa gggcatggcc tgtgtgggccc gcaccaagga atgtaccatc 1320
gtcccgctcca accactacgg acccatcccc gggatccccg tgggcacat gtggcggttc 1380
cgagtcacag tcagcgagtc ggggtgtccat cgccccacg tggctggcat acacggccgg 1440
agcaacgacg gagcgtatc cctagtctcg cgggggggct atgaggatga cgtggaccat 1500
gggaatTTTT tcacatacac gggtagtggt ggtcgagatc tttccggcaa caaggaggacc 1560
gcggaacagt cttgtgatca gaaactcacc aacaccaaca gggcgctggc tctcaactgc 1620
tttgtccca tcaatgacca agaaggggccc gagggcaagg actggcggtc ggggaagccg 1680
gtcaggggtg tgcgcaatgt caaggggtggc aagaatagca agtacgcccc cgctgagggc 1740
aaccgctacg atggcatcta caaggttgtg aaatactggc ccgagaaggg gaagtcggg 1800
tttctcgtgt ggcgtacct tctgcggagg gacgatgatg agcctggccc ttggacgaag 1860
gaggggaagg accgatcaa gaagctgggg ctgacctatg agtatccaga aggtacctg 1920
gaagccctgt ccaaccgaga gcgagagaag gagaacagca agagggagga ggaggagcag 1980
caggaggggg gcttcgcgtc ccccaggacg ggcaagggca agtggaaagc gaagtcggca 2040
ggaggtggcc cgagcagggc cgggtccccg cgccggacat ccaagaaaac caaggtggag 2100
ccctacagtc tcacggccca gcagagcagc ctcatcagag aggacaagag caacgccaa 2160
ctgtggaatg aggtcctggc gtcactcaag gaccggccgg cgagcggcag cccgttccag 2220
ttgttcctga gtaaaagtga ggagacgttc cagtgtatct gctgtcagga gctggtgttc 2280
cggcccatca cgaccgtgtg ccagcacaac gtgtgcaagg actgcctgga cagatccttt 2340
cgggcacagg tggtcagctg ccctgcctgc cgctacgacc tgggccgcag ctatgccatg 2400
caggtgaacc agcctctgca gaccgtctc aaccagctct tccccggcta cgggaatggc 2460
cggtgatctc caagcacttc tcgacaggcg ttttctgtaa aacgtgtcgg agggctcggt 2520
catcggcact gatTTTTgtt ttagtgggct taacttaaac aggtagtgtt tctccggtt 2580
cctaaaaagg tttgtcttcc ttttttttta tttttatatt tcaaatctat acattttcag 2640
gaatttatgt attctggcta aaagttggac ttctcagtat tgtgttttagt tctttgaaaa 2700

```

cataaaagcc tgcaatttct cgacaaaaca acacaagatt ttttaaagat ggaatcagaa 2760
actacgtggt gtggaggctg ttgatgttct tgggtgtcaag ttctcagaag ttgctgccac 2820
caactcttta agaaggcgac aggatcagtc cttctctagg gttctggccc ccaaggtcag 2880
agcaagcatc ttcctgacag cattttgtca tctaaagtcc agtgacatgg ttccccgtgg 2940
tggcccgtgg cagcccgtgg catggcgtgg ctcagctgtc tgttgaagtt gttgcaagga 3000
aaagaggaaa catctcgggc ctagttcaaa cctttgcctc aaagccatcc cccaccagac 3060
tgcttagcgt ctgagatccg cgtgaaaagt cctctgcccc cgagagcagg gagttggggc 3120
cacgcagaaa tggcctcaag gggactctgc tccacgtggg gccaggcgtg tgactgacgc 3180
tgtccgacga aggcgggccac ggacggacgc cagcacacga agtcacgtgc aagtgccttt 3240
gattcgttcc ttctttctaa agacgacagt ctttgttgtt agcactgaat tattgaaaat 3300
gtcaaccaga ttctagaaac tgcggtcatc cagttcttcc tgacaccgga tgggtgcttg 3360
ggaaccgttt gagccttata gatcatttac attcaatttt tttaactcag caagtgagaa 3420
cttacaagag gggttttttt taattttttt ttctcttaat gaacacattt tctaaatgaa 3480
ttttttttgt agttactgta tatgtaccaa gaaagatata acgttagggg ttggttgttt 3540
ttgtttttgt attttttttc ttttgaaagg gtttgttaat ttttctaatt ttaccaaagt 3600
ttgcagccta tacctcaata aaacagggat attttaaatc acatacctgc agacaaactg 3660
gagcaatgtt atttttaaaag gggttttttc acctccttat tcttagatta ttaatgtatt 3720
aggggaagaat gagacaattt tgtgtaggct ttttctaaag tccagtactt tgtccagatt 3780
ttagattctc agaataaatg tttttcacag attgaaaaaa aaaaaaaaa 3828

```

<210> 4

<211> 793

<212> PRT

<213> Homo sapiens

<220>

<223> nuclear zinc finger protein 95 (NP95)

<400> 4

```

Met Trp Ile Gln Val Arg Thr Met Asp Gly Arg Gln Thr His Thr Val
  1               5               10               15

```

```

Asp Ser Leu Ser Arg Leu Thr Lys Val Glu Glu Leu Arg Arg Lys Ile
      20               25               30

```

```

Gln Glu Leu Phe His Val Glu Pro Gly Leu Gln Arg Leu Phe Tyr Arg
      35               40               45

```

```

Gly Lys Gln Met Glu Asp Gly His Thr Leu Phe Asp Tyr Glu Val Arg
      50               55               60

```

```

Leu Asn Asp Thr Ile Gln Leu Leu Val Arg Gln Ser Leu Val Leu Pro
      65               70               75               80

```

```

His Ser Thr Lys Glu Arg Asp Ser Glu Leu Ser Asp Thr Asp Ser Gly
      85               90               95

```

```

Cys Cys Leu Gly Gln Ser Glu Ser Asp Lys Ser Ser Thr His Gly Glu
      100              105              110

```

```

Ala Ala Ala Glu Thr Asp Ser Arg Pro Ala Asp Glu Asp Met Trp Asp
      115              120              125

```

```

Glu Thr Glu Leu Gly Leu Tyr Lys Val Asn Glu Tyr Val Asp Ala Arg
      130              135              140

```

```

Asp Thr Asn Met Gly Ala Trp Phe Glu Ala Gln Val Val Arg Val Thr
      145              150              155              160

```

Arg	Lys	Ala	Pro	Ser	Arg	Asp	Glu	Pro	Cys	Ser	Ser	Thr	Ser	Arg	Pro	165	170	175
Ala	Leu	Glu	Glu	Asp	Val	Ile	Tyr	His	Val	Lys	Tyr	Asp	Asp	Tyr	Pro	180	185	190
Glu	Asn	Gly	Val	Val	Gln	Met	Asn	Ser	Arg	Asp	Val	Arg	Ala	Arg	Ala	195	200	205
Arg	Thr	Ile	Ile	Lys	Trp	Gln	Asp	Leu	Glu	Val	Gly	Gln	Val	Val	Met	210	215	220
Leu	Asn	Tyr	Asn	Pro	Asp	Asn	Pro	Lys	Glu	Arg	Gly	Phe	Trp	Tyr	Asp	225	230	235
Ala	Glu	Ile	Ser	Arg	Lys	Arg	Glu	Thr	Arg	Thr	Ala	Arg	Glu	Leu	Tyr	245	250	255
Ala	Asn	Val	Val	Leu	Gly	Asp	Asp	Ser	Leu	Asn	Asp	Cys	Arg	Ile	Ile	260	265	270
Phe	Val	Asp	Glu	Val	Phe	Lys	Ile	Glu	Arg	Pro	Gly	Glu	Gly	Ser	Pro	275	280	285
Met	Val	Asp	Asn	Pro	Met	Arg	Arg	Lys	Ser	Gly	Pro	Ser	Cys	Lys	His	290	295	300
Cys	Lys	Asp	Asp	Val	Asn	Arg	Leu	Cys	Arg	Val	Cys	Ala	Cys	His	Leu	305	310	315
Cys	Gly	Gly	Arg	Gln	Asp	Pro	Asp	Lys	Gln	Leu	Met	Cys	Asp	Glu	Cys	325	330	335
Asp	Met	Ala	Phe	His	Ile	Tyr	Cys	Leu	Asp	Pro	Pro	Leu	Ser	Ser	Val	340	345	350
Pro	Ser	Glu	Asp	Glu	Trp	Tyr	Cys	Pro	Glu	Cys	Arg	Asn	Asp	Ala	Ser	355	360	365
Glu	Val	Val	Leu	Ala	Gly	Glu	Arg	Leu	Arg	Glu	Ser	Lys	Lys	Lys	Ala	370	375	380
Lys	Met	Ala	Ser	Ala	Thr	Ser	Ser	Ser	Gln	Arg	Asp	Trp	Gly	Lys	Gly	385	390	395
Met	Ala	Cys	Val	Gly	Arg	Thr	Lys	Glu	Cys	Thr	Ile	Val	Pro	Ser	Asn	405	410	415
His	Tyr	Gly	Pro	Ile	Pro	Gly	Ile	Pro	Val	Gly	Thr	Met	Trp	Arg	Phe	420	425	430
Arg	Val	Gln	Val	Ser	Glu	Ser	Gly	Val	His	Arg	Pro	His	Val	Ala	Gly	435	440	445
Ile	His	Gly	Arg	Ser	Asn	Asp	Gly	Ala	Tyr	Ser	Leu	Val	Leu	Ala	Gly	450	455	460
Gly	Tyr	Glu	Asp	Asp	Val	Asp	His	Gly	Asn	Phe	Phe	Thr	Tyr	Thr	Gly	465	470	475
																		480

Ser Gly Gly Arg Asp Leu Ser Gly Asn Lys Arg Thr Ala Glu Gln Ser
 485 490 495
 Cys Asp Gln Lys Leu Thr Asn Thr Asn Arg Ala Leu Ala Leu Asn Cys
 500 505 510
 Phe Ala Pro Ile Asn Asp Gln Glu Gly Ala Glu Ala Lys Asp Trp Arg
 515 520 525
 Ser Gly Lys Pro Val Arg Val Val Arg Asn Val Lys Gly Gly Lys Asn
 530 535 540
 Ser Lys Tyr Ala Pro Ala Glu Gly Asn Arg Tyr Asp Gly Ile Tyr Lys
 545 550 555 560
 Val Val Lys Tyr Trp Pro Glu Lys Gly Lys Ser Gly Phe Leu Val Trp
 565 570 575
 Arg Tyr Leu Leu Arg Arg Asp Asp Asp Glu Pro Gly Pro Trp Thr Lys
 580 585 590
 Glu Gly Lys Asp Arg Ile Lys Lys Leu Gly Leu Thr Met Gln Tyr Pro
 595 600 605
 Glu Gly Tyr Leu Glu Ala Leu Ala Asn Arg Glu Arg Glu Lys Glu Asn
 610 615 620
 Ser Lys Arg Glu Glu Glu Glu Gln Gln Glu Gly Gly Phe Ala Ser Pro
 625 630 635 640
 Arg Thr Gly Lys Gly Lys Trp Lys Arg Lys Ser Ala Gly Gly Gly Pro
 645 650 655
 Ser Arg Ala Gly Ser Pro Arg Arg Thr Ser Lys Lys Thr Lys Val Glu
 660 665 670
 Pro Tyr Ser Leu Thr Ala Gln Gln Ser Ser Leu Ile Arg Glu Asp Lys
 675 680 685
 Ser Asn Ala Lys Leu Trp Asn Glu Val Leu Ala Ser Leu Lys Asp Arg
 690 695 700
 Pro Ala Ser Gly Ser Pro Phe Gln Leu Phe Leu Ser Lys Val Glu Glu
 705 710 715 720
 Thr Phe Gln Cys Ile Cys Cys Gln Glu Leu Val Phe Arg Pro Ile Thr
 725 730 735
 Thr Val Cys Gln His Asn Val Cys Lys Asp Cys Leu Asp Arg Ser Phe
 740 745 750
 Arg Ala Gln Val Phe Ser Cys Pro Ala Cys Arg Tyr Asp Leu Gly Arg
 755 760 765
 Ser Tyr Ala Met Gln Val Asn Gln Pro Leu Gln Thr Val Leu Asn Gln
 770 775 780
 Leu Phe Pro Gly Tyr Gly Asn Gly Arg
 785 790

<210> 5
 <211> 5503
 <212> DNA
 <213> Homo sapiens

<220>
 <223> Fanconi anemia, complementation group A protein
 (FANCA) (aldehyde dehydrogenase activity) cDNA

<400> 5
 agccgcgcgc ggggctgtag ggcgcaaggc catgtccgac tcgtgggtcc cgaactccgc 60
 ctcgggccag gaccagggg gccgcggag gccctgggag gagctgctgg cggaagggt 120
 caagagggaa aaatataatc ctgaaagggc acagaaatta aaggaatcag ctgtgcgct 180
 cctgcgaagc catcaggacc tgaatgccct tttgcttgag gtagaagggtc cactgtgtaa 240
 aaaattgtct ctacgcaaag tgattgactg tgacagttct gaggcctatg ctaatcattc 300
 tagttcattt ataggctctg ctttgaggga tcaagcctca aggctggggg ttcccgtggg 360
 tattctctca gccgggatgg ttgcctctag cgtgggacag atctgcacgg ctccagcgga 420
 gaccagtcac cctgtgctgc tgactgtgga gcagagaaag aagctgtctt ccctgttaga 480
 gtttgtctag tatttattgg cacacagtat gttctcccggt ctttccctct gtcaagaatt 540
 atggaaaata cagagttctt tgttgcttga agcgggtgtg catcttcacg tacaaggcat 600
 tgtgagcctg caagagctgc tggaaagcca tcccgcacatg catgctgtgg gatcgtgggt 660
 cttcaggaat ctgtgctgcc tttgtgaaca gatggaagca tcctgccagc atgctgacgt 720
 cgccaggggc atgctttctg attttgttca aatgtttgtt ttgaggggat ttcagaaaaa 780
 ctcatatctg agaagaactg tggagcctga aaaaatgccg caggtcacgg ttgatgtact 840
 gcagagaatg ctgatttttg cacttgacgc tttggctgct ggagtacagg aggagtcctc 900
 cactcacaag atcgtgaggt gctgggttcgg agtggttcagt ggacacacgc ttggcagtg 960
 aatttccaca gatcctctga agaggttctt cagtcatacc ctgactcaga tactcactca 1020
 cagccctgtg ctgaaagcat ctgatgctgt tcagatgcag agagagtgga gctttgcgcg 1080
 gacacaccct ctgctcacct cactgtaccg caggctcttt gtgatgctga gtgcagagga 1140
 gttggttggt catttgcaag aagtcttgga accgcaggag gttcactggc agagagtgtc 1200
 ctctttgtg tctgcctggg ttgtctgctt tccagaagcg cagcagctgc ttgaagactg 1260
 ggtggcgctg ttgatggccc aggcattcga gagctgccag ctggacagca tggctactgc 1320
 gttcctggtt gtgcgccagg cagcactgga gggccctctt gcgttctgt catatgcaga 1380
 ctggttcaag gcctcctttg ggagcacacg aggtaccat ggctgcagca agaaggcct 1440
 ggtcttctctg tttacgttct tgcagaact cgtgcctttt gagtctccc ggtacctgca 1500
 ggtgcacatt ctccaccac ccctggttcc cagcaagtac cgctccctcc tcacagacta 1560
 catctcattg gccaaagacac ggctggccga cctcaagggt tctatagaaa acatgggact 1620
 ctacaggatg ttgtcatcag ctggggacat tactgagccc cacagccaag ctcttcagga 1680
 tgttgaaaaa gccatcatgg tgtttgagca tacggggaac atcccagtca ccgtcatgga 1740
 ggccagcata tccaggaggc cttactacgt gtcccacttc ctcccgcgcc tgcacacac 1800
 tgcagtgtc cccaaagtcc ctgactcccg tgtggcggtt atagagtctc tgaagagagc 1860
 agataaaatc ccccatctc tgtactccac ctactgccag gcctgctctg ctgctgaaga 1920
 gaagccagaa gatgcagccc tgggagttag ggcagaacct aactctgctg aggagccct 1980
 gggacagctc acagctgcac tgggagagct gagagcctcc atgacagacc ccagccagcg 2040
 tgatgttata tcggcacagg tggcagtgat ttctgaaaga ctgagggctg tcctgggcca 2100
 caatgaggat gacagcagcg ttgagatata aaagattcag ctacagcatca acacgccag 2160
 actggagcca cgggaacaca ttgctgtgga cctcctgctg acgtctttct gtcagaacct 2220
 gatggctgcc tccagtgtcg ctcccccgga gaggcagggt ccctgggctg ccctctctct 2280
 gaggaccatg tgtggacgtg tgctccctgc agtgctcacc cggctctgcc agtgctccg 2340
 tcaccagggc ccgagcctga gtgccccaca tgtgctgggg ttggctgccc tggcctgca 2400
 cctgggtgag tccaggtctg cgctcccaga ggtggatgtg ggtcctctct cactggtgc 2460
 tggccttct gtccctgcgc tctttgacag cctcctgacc tgtaggacga gggattcctt 2520
 gttcttctgc ctgaaatttt gtacagcagc aatttcttac tctctctgca agtttctct 2580
 ccagtcacga gatactttgt gcagctgctt atctccaggc cttattaaaa agtttcagtt 2640
 cctcatgttc agattgttct cagaggcccc acagcctctt tctgaggagg acgtagccag 2700
 cctttcctgg agacccttgc accttcttct tgcagactgg cagagagctg ccctctctct 2760
 ctggacacac agaacccttc gagaggtgtt gaaagaggaa gatgttact taactacca 2820
 agactggtta cacctggagc tggaaattca acctgaagct gatgctcttt cagatactga 2880
 acggcaggac ttccaccagt gggcgatcca tgagcacttt ctccctgagt cctcggtctc 2940
 agggggctgt gacggagacc tgcaggctgc gtgtaccatt cttgtcaacg cactgatgga 3000
 tttccaccaa agctcaagga gttatgacca ctcaaaaaa tctgatttgg tctttggtgg 3060

```

ccgcacagga aatgaggata ttatttccag attgcaggag atggtagctg acctggagct 3120
gcagcaagac ctcatagtgc ctctcggcca cacccttcc caggagcact tcctccttga 3180
gattttccgc agacggctcc aggcctgtgac aagcgggtgg agcgtggctg ccagccttca 3240
gagacagagg gagctgctaa tgtacaaacg gatcctctc cgcttgcctt cgtctgtcct 3300
ctgcggcagc agcttccagg cagaacagcc catcactgcc agatgcgagc agttcttcca 3360
cttgggtcaac tctgagatga gaaacttctg ctcccacgga ggtgccctga cacaggacat 3420
cactgcccac ttcttcaggg gcctcctgaa cgctgtctg cggagcagag acccctccct 3480
gatggtcgac ttcatactgg ccaagtgcc aacgaaatgc cccttaattt tgacctctgc 3540
tctgggtgtg tggccgagcc tggagcctgt gctgctctgc cgggtggagga gacactgcc 3600
gagcccgtg ccccggaac tgcagaagct acaagaaggc cggcagtttg ccagcgattt 3660
cctctcccct gaggtgcct cccagcacc caaccggac tggctctcag ctgctgcact 3720
gcactttgag attcaacaag tcaggaaga aaacatcagg aagcagctaa agaagctgga 3780
ctgcgagaga gaggagctat tggtttctt tttcttctc tccttgatgg gcctgctgtc 3840
gtcacatctg acctcaaata gcaccacaga cctgccaaag gctttccacg tttgtgcagc 3900
aatcctcgag tgttttagaga agaggaagat atcctggctg gcaactcttc agttgacaga 3960
gagtgcctc aggtggggc ggctcctct cctgtgtggc ccgatcagc acaccaggct 4020
gctgcctttc gctttttaca gtcttctct ctacttccat gaagacgagg ccatcaggga 4080
agaggccttc ctgcatgttg ctgtggacat gtacttgaag ctggtccagc tcttcgtggc 4140
tggggatata agcacagttt cacctccagc tggcaggagc ctggagctca agggtcaggg 4200
caaccccggt gaactgataa caaaagctcg tcttttctg ctgcagttaa tacctcgggt 4260
ccgaaaaaag agcttctcac acgtggcaga gctgctggct gatcgtgggg actgcgaccc 4320
agaggtgagc gccgcctcc agagcagaca gcaggctgcc cctgacgctg acctgtccca 4380
ggagcctcat ctcttctgac gggacctgcc actgcacacc agcccagctc ccgtgtaaat 4440
aatttattac aagcataaca tggagctctt gttgcaactaa aaagtggatt acaaattctc 4500
tcgactgctt tagtggggaa aggaatcaat tatttatgaa ctgtccggcc ccgagtcact 4560
cagcgtttgc gggaaaataa accactgggt ccagagcaga ggaaggctac ttgagccgga 4620
caccaagccc gcctccagca ccaagggcgg gcagcacct ccgacctcc catgcggtg 4680
cacacgaagg gtgaggctga cacagccact gcggagtcca ggctgctaga ggtgctcatc 4740
ctcactgcc tcctcaggtg ggttcgggt tcaccgctg gccctctgtg gtcacagagg 4800
ggctcgggtg ccaggtgggt ggttcgcct ccaggggcag ggccttgtcc tgggtctgtg 4860
tcagcgggtg caccatggac atgtgtacat tgaggttggt ggccttctca aaccgccggc 4920
cacactgggt acaggcaaag tccagctcag tctcagcctt gtgtttgggt atgtggtact 4980
tgagggatgc ccgctgcctg cactggaacc cacagacctc acacctgggg gacagaggca 5040
gataagaagg tgcgaggcca cagccctggg agggggtcct gactcacact tactgcaaag 5100
gcttggtctc cgaatgtcgc atttggtgga cgagaagggt cttccgctgc ttgaaggttt 5160
gtccacattc gtcacagata tagttccgca cctctgagag gggagagtcc agtgagtcca 5220
ggccctgat gctccaacct cccgggggga cgacgatgac aatgtgaaac catcacagct 5280
gggaagacat ttctgcacat ggttcacat gcagtgggca caagcaagg gcctatgagg 5340
gcctcgttta ttaagatctt taaactgctt tataactgt cagctggctt catcagctgt 5400
gtgcatttca ggatggtttt taaagaaacc tcagaaagct atttccttaa aaaaaaaaaa 5460
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaa 5503

```

<210> 6

<211> 1455

J <212> PRT

<213> Homo sapiens

<220>

<223> Fanconi anemia, complementation group A protein
(FANCA) (aldehyde dehydrogenase activity)

<400> 6

Met Ser Asp Ser Trp Val Pro Asn Ser Ala Ser Gly Gln Asp Pro Gly
1 5 10 15

Gly Arg Arg Arg Ala Trp Ala Glu Leu Leu Ala Gly Arg Val Lys Arg
20 25 30

Glu Lys Tyr Asn Pro Glu Arg Ala Gln Lys Leu Lys Glu Ser Ala Val
35 40 45

Arg	Leu	Leu	Arg	Ser	His	Gln	Asp	Leu	Asn	Ala	Leu	Leu	Leu	Glu	Val	50	55	60
Glu	Gly	Pro	Leu	Cys	Lys	Lys	Leu	Ser	Leu	Ser	Lys	Val	Ile	Asp	Cys	65	70	75
Asp	Ser	Ser	Glu	Ala	Tyr	Ala	Asn	His	Ser	Ser	Ser	Phe	Ile	Gly	Ser	85	90	95
Ala	Leu	Gln	Asp	Gln	Ala	Ser	Arg	Leu	Gly	Val	Pro	Val	Gly	Ile	Leu	100	105	110
Ser	Ala	Gly	Met	Val	Ala	Ser	Ser	Val	Gly	Gln	Ile	Cys	Thr	Ala	Pro	115	120	125
Ala	Glu	Thr	Ser	His	Pro	Val	Leu	Leu	Thr	Val	Glu	Gln	Arg	Lys	Lys	130	135	140
Leu	Ser	Ser	Leu	Leu	Glu	Phe	Ala	Gln	Tyr	Leu	Leu	Ala	His	Ser	Met	145	150	155
Phe	Ser	Arg	Leu	Ser	Phe	Cys	Gln	Glu	Leu	Trp	Lys	Ile	Gln	Ser	Ser	165	170	175
Leu	Leu	Leu	Glu	Ala	Val	Trp	His	Leu	His	Val	Gln	Gly	Ile	Val	Ser	180	185	190
Leu	Gln	Glu	Leu	Leu	Glu	Ser	His	Pro	Asp	Met	His	Ala	Val	Gly	Ser	195	200	205
Trp	Leu	Phe	Arg	Asn	Leu	Cys	Cys	Leu	Cys	Glu	Gln	Met	Glu	Ala	Ser	210	215	220
Cys	Gln	His	Ala	Asp	Val	Ala	Arg	Ala	Met	Leu	Ser	Asp	Phe	Val	Gln	225	230	235
Met	Phe	Val	Leu	Arg	Gly	Phe	Gln	Lys	Asn	Ser	Asp	Leu	Arg	Arg	Thr	245	250	255
Val	Glu	Pro	Glu	Lys	Met	Pro	Gln	Val	Thr	Val	Asp	Val	Leu	Gln	Arg	260	265	270
Met	Leu	Ile	Phe	Ala	Leu	Asp	Ala	Leu	Ala	Ala	Gly	Val	Gln	Glu	Glu	275	280	285
Ser	Ser	Thr	His	Lys	Ile	Val	Arg	Cys	Trp	Phe	Gly	Val	Phe	Ser	Gly	290	295	300
His	Thr	Leu	Gly	Ser	Val	Ile	Ser	Thr	Asp	Pro	Leu	Lys	Arg	Phe	Phe	305	310	315
Ser	His	Thr	Leu	Thr	Gln	Ile	Leu	Thr	His	Ser	Pro	Val	Leu	Lys	Ala	325	330	335
Ser	Asp	Ala	Val	Gln	Met	Gln	Arg	Glu	Trp	Ser	Phe	Ala	Arg	Thr	His	340	345	350
Pro	Leu	Leu	Thr	Ser	Leu	Tyr	Arg	Arg	Leu	Phe	Val	Met	Leu	Ser	Ala	355	360	365

Glu 370	Leu	Val	Gly	His	Leu 375	Gln	Glu	Val	Leu	Glu 380	Thr	Gln	Glu	Val	
His 385	Trp	Gln	Arg	Val	Leu 390	Ser	Phe	Val	Ser	Ala 395	Leu	Val	Val	Cys	Phe 400
Pro	Glu	Ala	Gln	Gln 405	Leu	Leu	Glu	Asp	Trp 410	Val	Ala	Arg	Leu	Met	Ala 415
Gln	Ala	Phe	Glu	Ser	Cys	Gln	Leu	Asp 425	Ser	Met	Val	Thr	Ala 430	Phe	Leu
Val	Val	Arg 435	Gln	Ala	Ala	Leu	Glu 440	Gly	Pro	Ser	Ala	Phe 445	Leu	Ser	Tyr
Ala	Asp 450	Trp	Phe	Lys	Ala	Ser 455	Phe	Gly	Ser	Thr	Arg 460	Gly	Tyr	His	Gly
Cys 465	Ser	Lys	Lys	Ala	Leu 470	Val	Phe	Leu	Phe	Thr 475	Phe	Leu	Ser	Glu	Leu 480
Val	Pro	Phe	Glu	Ser 485	Pro	Arg	Tyr	Leu	Gln 490	Val	His	Ile	Leu	His 495	Pro
Pro	Leu	Val	Pro 500	Ser	Lys	Tyr	Arg	Ser 505	Leu	Leu	Thr	Asp	Tyr 510	Ile	Ser
Leu	Ala	Lys 515	Thr	Arg	Leu	Ala	Asp 520	Leu	Lys	Val	Ser	Ile 525	Glu	Asn	Met
Gly	Leu 530	Tyr	Glu	Asp	Leu	Ser 535	Ser	Ala	Gly	Asp	Ile 540	Thr	Glu	Pro	His
Ser 545	Gln	Ala	Leu	Gln 550	Asp	Val	Glu	Lys	Ala	Ile 555	Met	Val	Phe	Glu	His 560
Thr	Gly	Asn	Ile	Pro 565	Val	Thr	Val	Met	Glu 570	Ala	Ser	Ile	Phe	Arg 575	Arg
Pro	Tyr	Tyr	Val 580	Ser	His	Phe	Leu	Pro 585	Ala	Leu	Leu	Thr	Pro 590	Arg	Val
Leu	Pro	Lys 595	Val	Pro	Asp	Ser	Arg 600	Val	Ala	Phe	Ile	Glu 605	Ser	Leu	Lys
Arg	Ala 610	Asp	Lys	Ile	Pro	Pro 615	Ser	Leu	Tyr	Ser	Thr 620	Tyr	Cys	Gln	Ala
Cys 625	Ser	Ala	Ala	Glu	Glu 630	Lys	Pro	Glu	Asp	Ala 635	Ala	Leu	Gly	Val	Arg 640
Ala	Glu	Pro	Asn	Ser 645	Ala	Glu	Glu	Pro	Leu 650	Gly	Gln	Leu	Thr	Ala 655	Ala
Leu	Gly	Glu	Leu 660	Arg	Ala	Ser	Met	Thr 665	Asp	Pro	Ser	Gln	Arg 670	Asp	Val
Ile	Ser	Ala 675	Gln	Val	Ala	Val	Ile 680	Ser	Glu	Arg	Leu	Arg 685	Ala	Val	Leu

Gly	His	Asn	Glu	Asp	Asp	Ser	Ser	Val	Glu	Ile	Ser	Lys	Ile	Gln	Leu	690	695	700
Ser	Ile	Asn	Thr	Pro	Arg	Leu	Glu	Pro	Arg	Glu	His	Ile	Ala	Val	Asp	705	710	715
Leu	Leu	Leu	Thr	Ser	Phe	Cys	Gln	Asn	Leu	Met	Ala	Ala	Ser	Ser	Val	725	730	735
Ala	Pro	Pro	Glu	Arg	Gln	Gly	Pro	Trp	Ala	Ala	Leu	Phe	Val	Arg	Thr	740	745	750
Met	Cys	Gly	Arg	Val	Leu	Pro	Ala	Val	Leu	Thr	Arg	Leu	Cys	Gln	Leu	755	760	765
Leu	Arg	His	Gln	Gly	Pro	Ser	Leu	Ser	Ala	Pro	His	Val	Leu	Gly	Leu	770	775	780
Ala	Ala	Leu	Ala	Val	His	Leu	Gly	Glu	Ser	Arg	Ser	Ala	Leu	Pro	Glu	785	790	795
Val	Asp	Val	Gly	Pro	Pro	Ala	Pro	Gly	Ala	Gly	Leu	Pro	Val	Pro	Ala	805	810	815
Leu	Phe	Asp	Ser	Leu	Leu	Thr	Cys	Arg	Thr	Arg	Asp	Ser	Leu	Phe	Phe	820	825	830
Cys	Leu	Lys	Phe	Cys	Thr	Ala	Ala	Ile	Ser	Tyr	Ser	Leu	Cys	Lys	Phe	835	840	845
Ser	Ser	Gln	Ser	Arg	Asp	Thr	Leu	Cys	Ser	Cys	Leu	Ser	Pro	Gly	Leu	850	855	860
Ile	Lys	Lys	Phe	Gln	Phe	Leu	Met	Phe	Arg	Leu	Phe	Ser	Glu	Ala	Arg	865	870	875
Gln	Pro	Leu	Ser	Glu	Glu	Asp	Val	Ala	Ser	Leu	Ser	Trp	Arg	Pro	Leu	885	890	895
His	Leu	Pro	Ser	Ala	Asp	Trp	Gln	Arg	Ala	Ala	Leu	Ser	Leu	Trp	Thr	900	905	910
His	Arg	Thr	Phe	Arg	Glu	Val	Leu	Lys	Glu	Glu	Asp	Val	His	Leu	Thr	915	920	925
Tyr	Gln	Asp	Trp	Leu	His	Leu	Glu	Leu	Glu	Ile	Gln	Pro	Glu	Ala	Asp	930	935	940
Ala	Leu	Ser	Asp	Thr	Glu	Arg	Gln	Asp	Phe	His	Gln	Trp	Ala	Ile	His	945	950	955
Glu	His	Phe	Leu	Pro	Glu	Ser	Ser	Ala	Ser	Gly	Gly	Cys	Asp	Gly	Asp	965	970	975
Leu	Gln	Ala	Ala	Cys	Thr	Ile	Leu	Val	Asn	Ala	Leu	Met	Asp	Phe	His	980	985	990
Gln	Ser	Ser	Arg	Ser	Tyr	Asp	His	Ser	Glu	Asn	Ser	Asp	Leu	Val	Phe	995	1000	1005

Gly Gly Arg Thr Gly Asn Glu Asp Ile Ile Ser Arg Leu Gln Glu Met
 1010 1015 1020
 Val Ala Asp Leu Glu Leu Gln Gln Asp Leu Ile Val Pro Leu Gly His
 1025 1030 1035 1040
 Thr Pro Ser Gln Glu His Phe Leu Phe Glu Ile Phe Arg Arg Arg Leu
 1045 1050 1055
 Gln Ala Leu Thr Ser Gly Trp Ser Val Ala Ala Ser Leu Gln Arg Gln
 1060 1065 1070
 Arg Glu Leu Leu Met Tyr Lys Arg Ile Leu Leu Arg Leu Pro Ser Ser
 1075 1080 1085
 Val Leu Cys Gly Ser Ser Phe Gln Ala Glu Gln Pro Ile Thr Ala Arg
 1090 1095 1100
 Cys Glu Gln Phe Phe His Leu Val Asn Ser Glu Met Arg Asn Phe Cys
 1105 1110 1115 1120
 Ser His Gly Gly Ala Leu Thr Gln Asp Ile Thr Ala His Phe Phe Arg
 1125 1130 1135
 Gly Leu Leu Asn Ala Cys Leu Arg Ser Arg Asp Pro Ser Leu Met Val
 1140 1145 1150
 Asp Phe Ile Leu Ala Lys Cys Gln Thr Lys Cys Pro Leu Ile Leu Thr
 1155 1160 1165
 Ser Ala Leu Val Trp Trp Pro Ser Leu Glu Pro Val Leu Leu Cys Arg
 1170 1175 1180
 Trp Arg Arg His Cys Gln Ser Pro Leu Pro Arg Glu Leu Gln Lys Leu
 1185 1190 1195 1200
 Gln Glu Gly Arg Gln Phe Ala Ser Asp Phe Leu Ser Pro Glu Ala Ala
 1205 1210 1215
 Ser Pro Ala Pro Asn Pro Asp Trp Leu Ser Ala Ala Ala Leu His Phe
 1220 1225 1230
 Ala Ile Gln Gln Val Arg Glu Glu Asn Ile Arg Lys Gln Leu Lys Lys
 1235 1240 1245
 Leu Asp Cys Glu Arg Glu Glu Leu Leu Val Phe Leu Phe Phe Phe Ser
 1250 1255 1260
 Leu Met Gly Leu Leu Ser Ser His Leu Thr Ser Asn Ser Thr Thr Asp
 1265 1270 1275 1280
 Leu Pro Lys Ala Phe His Val Cys Ala Ala Ile Leu Glu Cys Leu Glu
 1285 1290 1295
 Lys Arg Lys Ile Ser Trp Leu Ala Leu Phe Gln Leu Thr Glu Ser Asp
 1300 1305 1310
 Leu Arg Leu Gly Arg Leu Leu Leu Arg Val Ala Pro Asp Gln His Thr
 1315 1320 1325

Arg Leu Leu Pro Phe Ala Phe Tyr Ser Leu Leu Ser Tyr Phe His Glu
 1330 1335 1340

Asp Ala Ala Ile Arg Glu Glu Ala Phe Leu His Val Ala Val Asp Met
 1345 1350 1355 1360

Tyr Leu Lys Leu Val Gln Leu Phe Val Ala Gly Asp Thr Ser Thr Val
 1365 1370 1375

Ser Pro Pro Ala Gly Arg Ser Leu Glu Leu Lys Gly Gln Gly Asn Pro
 1380 1385 1390

Val Glu Leu Ile Thr Lys Ala Arg Leu Phe Leu Leu Gln Leu Ile Pro
 1395 1400 1405

Arg Cys Pro Lys Lys Ser Phe Ser His Val Ala Glu Leu Leu Ala Asp
 1410 1415 1420

Arg Gly Asp Cys Asp Pro Glu Val Ser Ala Ala Leu Gln Ser Arg Gln
 1425 1430 1435 1440

Gln Ala Ala Pro Asp Ala Asp Leu Ser Gln Glu Pro His Leu Phe
 1445 1450 1455

<210> 7

<211> 1378

<212> DNA

<213> Homo sapiens

<220>

<223> DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 9
 (DDX9) (RNA helicase A, nuclear DNA helicase II,
 leukophysin), transcript variant 2 cDNA

<400> 7

```
cattgctgct gctacctgct ttccagagcc ttctcatcaat gaaggaaagc ggctgggcta 60
tatccatcga aatttttgctg gaaacagatt ttctgatcac gtagcccttt tatcagtatt 120
ccaagcctgg gatgatgcta gaatgggtgg agaagaagca gagatacgtt tttgtgagca 180
caaaagactt aatatggcta cactaagaat gacctgggaa gccaaagtgc agctcaaaga 240
gatttttgatt aattctgggt ttccagaaga ttgtttgttg acacaagtgt ttactaacac 300
tggaccagat aataatttgg atgttggtat ctccctcctg gcctttgtag ccaagacatg 360
aagtacccat ctcccttctt tgtatttggg gaaaagattc gaactcgagc catctctgct 420
aaaggcatga ctttagtcac cccctgcag ttgcttctct ttgcctccaa gaaagtccaa 480
tctgatgggc agattgtgct tgtagatgac tggattaaac tgcaaataatc tcatgaagct 540
getgcctgta tcaactggtc cggggcagcc atggaggctt tggttgttga agtaaccaa 600
caacctgcta tcatcagcca gttggacccc gtaaataaac gtatgctgaa catgatccgt 660
cagatctcta gacctcagc tgcctggatc aaccttatga ttggcagtag acggtatgga 720
gatgggtccac gtcctcccaa gatggcccga tacgacaatg gaagcggata tagaagggga 780
ggttctagtt acagtgggtg aggctatggc ggtggctata gcagtggagg ctatggtagc 840
ggaggctatg gtggcagcgc caactccttt cgggcaggat atgggtgcagg tgttgggtgga 900
ggctatagag gagtttcccg aggtggcttt agaggcaact ctggaggaga ctacagaggg 960
cctagtggag gctacagagg atctggggga ttccagcgag gaggtggtag gggggcctat 1020
ggaactggct actttggaca gggaagagga ggtggcggct attaaaactt ggttatgtca 1080
gttcctgtgt gtagacagta aggaaaaaaa ggcattgctat gtgttacgtg ttttttccag 1140
tatgtttatt tgccaccaa aagtaaatgc attttcacc attctgtggt tcattgtagt 1200
ttaaggaaac caagcatata gatgcattag tgattttgtt tatattatgt aaaatataac 1260
gatctcttaa aaataccaca gtttgtattt tttctttaag gagtaaagat ttgcctttaa 1320
ataacttggt attttcctgg ctttcgttta atacaataga aaataaagta ttacaccg 1378
```

<210> 8
 <211> 235
 <212> PRT
 <213> Homo sapiens

<220>
 <223> DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 9
 (DDX9) (RNA helicase A, nuclear DNA helicase II,
 leukophysin), transcript variant 2

<400> 8
 Met Lys Tyr Pro Ser Pro Phe Phe Val Phe Gly Glu Lys Ile Arg Thr
 1 5 10 15
 Arg Ala Ile Ser Ala Lys Gly Met Thr Leu Val Thr Pro Leu Gln Leu
 20 25 30
 Leu Leu Phe Ala Ser Lys Lys Val Gln Ser Asp Gly Gln Ile Val Leu
 35 40 45
 Val Asp Asp Trp Ile Lys Leu Gln Ile Ser His Glu Ala Ala Ala Cys
 50 55 60
 Ile Thr Gly Leu Arg Ala Ala Met Glu Ala Leu Val Val Glu Val Thr
 65 70 75 80
 Lys Gln Pro Ala Ile Ile Ser Gln Leu Asp Pro Val Asn Glu Arg Met
 85 90 95
 Leu Asn Met Ile Arg Gln Ile Ser Arg Pro Ser Ala Ala Gly Ile Asn
 100 105 110
 Leu Met Ile Gly Ser Thr Arg Tyr Gly Asp Gly Pro Arg Pro Pro Lys
 115 120 125
 Met Ala Arg Tyr Asp Asn Gly Ser Gly Tyr Arg Arg Gly Gly Ser Ser
 130 135 140
 Tyr Ser Gly Gly Gly Tyr Gly Gly Gly Tyr Ser Ser Gly Gly Tyr Gly
 145 150 155 160
 Ser Gly Gly Tyr Gly Gly Ser Ala Asn Ser Phe Arg Ala Gly Tyr Gly
 165 170 175
 Ala Gly Val Gly Gly Gly Tyr Arg Gly Val Ser Arg Gly Gly Phe Arg
 180 185 190
 Gly Asn Ser Gly Gly Asp Tyr Arg Gly Pro Ser Gly Gly Tyr Arg Gly
 195 200 205
 Ser Gly Gly Phe Gln Arg Gly Gly Gly Arg Gly Ala Tyr Gly Thr Gly
 210 215 220
 Tyr Phe Gly Gln Gly Arg Gly Gly Gly Gly Tyr
 225 230 235

<210> 9
 <211> 4989
 <212> DNA
 <213> Homo sapiens

<220>
 <223> insulin-like growth factor 1 receptor (IGF1R)
 (cell surface kinase receptor) cDNA

<400> 9
 tttttttttt ttttgagaaa ggggaatttca tcccaaataa aaggaatgaa gtctgggtcc 60
 ggaggagggt ccccgacctc gctgtggggg ctctgttttc tctccgccgc gctctcgctc 120
 tggccgacga gtggagaaat ctgctgggcca ggcctcgaca tccgcaacga ctatcagcag 180
 ctgaagcgcc tggagaactg cacggtgate gaggggtacc tccacatcct gctcatctcc 240
 aaggccgagg actaccgcag ctaccgcttc cccaagctca cggtcattac cgagtacttg 300
 ctgctgtttc gagtggctgg cctcgagagc ctgggagacc tcttcccaa cctcacggtc 360
 atccgcggtt ggaaactctt ctacaactac gccctgggtc tcttcgagat gaccaatctc 420
 aaggatatgt ggctttacaa cctgaggaac attactcggg gggccatcag gattgagaaa 480
 aatgctgacc tctgtttacct ctccactgtg gactgggtccc tgatcctgga tgcgggtgtcc 540
 aataactaca ttgtggggaa taagccccc aaggaatgtg gggacctgtg tccagggacc 600
 atggaggaga agccgatgtg tgagaagacc accatcaaca atgagtacaa ctaccgctgc 660
 tggaccacaa accgctgcca gaaaatgtgc ccaagcacgt gtgggaagcg ggcgtgcacc 720
 gagaacaatg agtgctgcca ccccgagtgc ctgggcagct gcagcgcgcc tgacaacgac 780
 acggcctgtg tagcttgccg ccactactac tatgccggtg tctgtgtgcc tgccctgccc 840
 cccaacacct acaggtttga gggctggcgc tgtgtggacc gtgacttctg cgccaacatc 900
 ctcagcgccg agagcagcga tctcgagggg tttgtgatcc acgacggcga gtgcattgcag 960
 gagtgcacct cggtcttcac ccgcaacggc agccagagca tgtactgcat cccttgtgaa 1020
 ggtccttgcc cgaagggtctg tgaggaagaa aagaaaacaa agaccattga ttctgttact 1080
 tctgctcaga tgctccaagg atgcaccatc ttcaagggca atttgctcat taacatccga 1140
 cgggggaata acattgtctc agagctggag aacttcatgg ggctcatcga ggtggtgacg 1200
 ggctacgtga agatccgcca ttctcatgcc ttggtctcct tgtccttctc aaaaaacctt 1260
 cgctcatcc taggagagga gcagctagaa ggggaattact ccttctacgt cctcgacaac 1320
 cagaacttgc agcaactgtg ggactgggac caccgcaacc tgaccatcaa agcagggaaa 1380
 atgtactttg ctttcaatcc caaattatgt gtttccgaaa tttaccgcat ggaggaagtg 1440
 acggggccta aagggcgcca aagcaaaagg gacataaaca ccaggaacaa cggggagaga 1500
 gcctcctgtg aaagtgcagt cctgcatttc acctccacca ccacgtcgaa gaatcgcatac 1560
 atcataacct ggcaccggta ccggccccct gactacaggg atctcatcag cttcaccgtt 1620
 tactacaagg aagcaccctt taagaatgtc acagagtatg atgggcagga tgcttgcggc 1680
 tccaacagct ggaacatggt ggacgtggac ctcccgccca acaaggacgt ggagcccggc 1740
 atcttactac atgggctgaa gccctggact cagtacgccg ttacgtcaa ggctgtgacc 1800
 ctaccatgg tggagaacga ccatatccgt ggggccaaaga gtgagatctt gtacattcgc 1860
 accaatgctt cagttccttc cattcccttg gacgttcttt cagcatcgaa ctctcttctt 1920
 cagttaatcg tgaagtggaa ccctccctct ctgccccacg gcaacctgag ttactacatt 1980
 gtgcgtggtg agcggcagcc tcaggacggc tacctttacc ggcacaatta ctgctccaaa 2040
 gacaaaatcc ccatcaggaa gtatgccgac ggcaccatcg acattgagga ggtcacagag 2100
 aaccccaaga ctgaggtgtg tgggtggggag aaagggcctt gctgcgcctg ccccaaaact 2160
 gaagccgaga agcaggccga gaaggaggag gctgaatacc gcaaagtctt tgagaatttc 2220
 ctgcacaact ccatcttcgt gccagacct gaaaggaagc ggagagatgt catgcaagtg 2280
 gccaacacca ccatgtccag ccgaagcagg aacaccacgg ccgcagacac ctacaacatc 2340
 accgaccccg aagagctgga gacagagtac cctttctttg agagcagagt ggataacaag 2400
 gagagaactg tcatttctaa ccttcggcct ttacattgt accgcatcga tatccacagc 2460
 tgcaaccacg aggctgagaa gctgggctgc agcgctcca acttcgtctt tgcaaggact 2520
 atgcccgcag aaggagcaga tgacattcct gggccagtga cctgggagcc aaggcctgaa 2580
 aactccatct ttttaaagtg gccggaacct gagaatccca atggattgat tctaattgat 2640
 gaaataaaat acggatcaca agttgaggat cagcgagaaat gtgtgtccag acaggaatac 2700
 aggaagtatg gaggggcca gctaaaccgg cttaaaccgg ggaactacac agcccgatt 2760
 caggccacat ctctctctgg gaatgggtcg tggacagatc ctgtgttctt ctatgtccag 2820
 gccaaaacag gatatgaaaa cttcatccat ctgatcatcg ctctgcccgt cgctgtcctg 2880
 ttgatcgtgg gaggttggt gattatgctg tacgtcttcc atagaaagag aaataacagc 2940
 aggctgggga atggagtgt gtatgcctct gtgaaccgg agtacttcag cgctgtgtat 3000
 gtgtacgttc ctgatgagt ggaggtggct cgggagaaga tcaccatgag ccgggaactt 3060
 gggcaggggt cgtttgggat ggtctatgaa aggggtgtgt caagcatgag tgagaggatt 3120
 cctgaaacca gagtggccat taaaacagtg aacgaggccg caagcatgag tgagaggatt 3180
 gagtttctca acgaagcttc tgtgatgaag gagttcaatt gtcaccatgt ggtgcgattg 3240
 ctgggtgtgg tgtcccaagg ccagccaaca ctggtcatca tggaaactgat gacacggggc 3300
 gatctcaaaa gttatctccg gtctctgagg ccagaaatgg agaataatcc agtcctagca 3360

```

cctccaagcc tgagcaagat gattcagatg gccggagaga ttgcagacgg catggcatac 3420
ctcaacgcca ataatgtcgt ccacagagac cttgctgccc ggaattgcat ggtagccgaa 3480
gatttcacag tcaaaatcgg agatttttgg atgacgcgag atatctatga gacagactat 3540
taccggaaaag gagggaaaagg gctgctgccc gtgcgctgga tgtctcctga gtccctcaag 3600
gatggagtct tcaccactta ctcggaagtc tggctcctcg gggtcgtcct ctgggagatc 3660
gccacactgg ccgagcagcc ctaccagggc ttgtccaacg agcaagtcct tcgcttcgtc 3720
atggagggcg gccttctgga caagccagac aactgtcctg acatgctgtt tgaactgatg 3780
cgcatgtgct ggaggtataa cccaagatg aggccttcct tcctggagat catcagcagc 3840
atcaaagagg agatggagcc tggcttcgg gaggtctcct tctactacag cgaggagaac 3900
aagctgcccc agccggagga gctggacctg gagccagaga acatggagag cgtccccctg 3960
gaccctcgg cctcctcgtc ctccctgcca ctgcccgaca gacactcagg acacaaggcc 4020
gagaacggcc ccggccctgg ggtgctggtc ctccgcgcca gcttcgacga gagacagcct 4080
tacgcccaca tgaacggggg ccgcaagaac gagcgggcct tgccgctgcc ccagtcttcg 4140
acctgctgat ccttggatcc tgaatctgtg caaacagtaa cgtgtgcgca cgcgcagcgg 4200
ggtggggggg gagagagagt tttaacaatc cattacaag cctcctgtac ctcagtggat 4260
cttcagttct gcccttgctg ccgcggggag acagcttctc tgcagtaaaa cacatttggg 4320
atgttccttt tttcaatatg caagcagctt tttattccct gcccaaacc ttaactgaca 4380
tgggccttta agaaccttaa tgacaacact taatagcaac agagcacttg agaaccagtc 4440
tcctcactct gtccctgtcc ttccctgttc tccctttctc tctcctctct gcttcataac 4500
ggaaaaataa ttgccacaag tccagctggg aagccctttt tatcagtttg aggaagtggc 4560
tgtccctgtg gccccatcca accactgtac acaccgcct gacaccgtgg gtcattacaa 4620
aaaaacacgt ggagatggaa atttttacct ttatctttca cctttctagg gacatgaaat 4680
ttacaaaagg ccacgttca tccaaggctg ttaccatttt aacgctgcct aattttgcca 4740
aaatcctgaa ctttctccct catcgcccgc gcgctgattc ctgctgtccg gaggcattgg 4800
tgagcatggc agctgggtgc tccatttgag agacacgctg gcgacacact ccgtccatcc 4860
gactgcccct gctgtgctgc tcaaggccac aggcacacag gtctcattgc ttctgactag 4920
attattattt gggggaactg gacacaatag gtctttctct cagtgaaggt ggggagaagc 4980
tgaaccggc 4989

```

<210> 10

<211> 1367

<212> PRT

<213> Homo sapiens

<220>

<223> insulin-like growth factor 1 receptor (IGF1R)
(cell surface kinase receptor)

<400> 10

```

Met Lys Ser Gly Ser Gly Gly Gly Ser Pro Thr Ser Leu Trp Gly Leu
  1                      5                      10          15

```

```

Leu Phe Leu Ser Ala Ala Leu Ser Leu Trp Pro Thr Ser Gly Glu Ile
          20                      25          30

```

```

Cys Gly Pro Gly Ile Asp Ile Arg Asn Asp Tyr Gln Gln Leu Lys Arg
          35                      40          45

```

```

Leu Glu Asn Cys Thr Val Ile Glu Gly Tyr Leu His Ile Leu Leu Ile
          50                      55          60

```

```

Ser Lys Ala Glu Asp Tyr Arg Ser Tyr Arg Phe Pro Lys Leu Thr Val
          65                      70          75          80

```

```

Ile Thr Glu Tyr Leu Leu Leu Phe Arg Val Ala Gly Leu Glu Ser Leu
          85                      90          95

```

```

Gly Asp Leu Phe Pro Asn Leu Thr Val Ile Arg Gly Trp Lys Leu Phe
          100                      105          110

```

Tyr	Asn	Tyr	Ala	Leu	Val	Ile	Phe	Glu	Met	Thr	Asn	Leu	Lys	Asp	Ile	115	120	125
Gly	Leu	Tyr	Asn	Leu	Arg	Asn	Ile	Thr	Arg	Gly	Ala	Ile	Arg	Ile	Glu	130	135	140
Lys	Asn	Ala	Asp	Leu	Cys	Tyr	Leu	Ser	Thr	Val	Asp	Trp	Ser	Leu	Ile	145	150	155
Leu	Asp	Ala	Val	Ser	Asn	Asn	Tyr	Ile	Val	Gly	Asn	Lys	Pro	Pro	Lys	165	170	175
Glu	Cys	Gly	Asp	Leu	Cys	Pro	Gly	Thr	Met	Glu	Glu	Lys	Pro	Met	Cys	180	185	190
Glu	Lys	Thr	Thr	Ile	Asn	Asn	Glu	Tyr	Asn	Tyr	Arg	Cys	Trp	Thr	Thr	195	200	205
Asn	Arg	Cys	Gln	Lys	Met	Cys	Pro	Ser	Thr	Cys	Gly	Lys	Arg	Ala	Cys	210	215	220
Thr	Glu	Asn	Asn	Glu	Cys	Cys	His	Pro	Glu	Cys	Leu	Gly	Ser	Cys	Ser	225	230	235
Ala	Pro	Asp	Asn	Asp	Thr	Ala	Cys	Val	Ala	Cys	Arg	His	Tyr	Tyr	Tyr	245	250	255
Ala	Gly	Val	Cys	Val	Pro	Ala	Cys	Pro	Pro	Asn	Thr	Tyr	Arg	Phe	Glu	260	265	270
Gly	Trp	Arg	Cys	Val	Asp	Arg	Asp	Phe	Cys	Ala	Asn	Ile	Leu	Ser	Ala	275	280	285
Glu	Ser	Ser	Asp	Ser	Glu	Gly	Phe	Val	Ile	His	Asp	Gly	Glu	Cys	Met	290	295	300
Gln	Glu	Cys	Pro	Ser	Gly	Phe	Ile	Arg	Asn	Gly	Ser	Gln	Ser	Met	Tyr	305	310	315
Cys	Ile	Pro	Cys	Glu	Gly	Pro	Cys	Pro	Lys	Val	Cys	Glu	Glu	Glu	Lys	325	330	335
Lys	Thr	Lys	Thr	Ile	Asp	Ser	Val	Thr	Ser	Ala	Gln	Met	Leu	Gln	Gly	340	345	350
Cys	Thr	Ile	Phe	Lys	Gly	Asn	Leu	Leu	Ile	Asn	Ile	Arg	Arg	Gly	Asn	355	360	365
Asn	Ile	Ala	Ser	Glu	Leu	Glu	Asn	Phe	Met	Gly	Leu	Ile	Glu	Val	Val	370	375	380
Thr	Gly	Tyr	Val	Lys	Ile	Arg	His	Ser	His	Ala	Leu	Val	Ser	Leu	Ser	385	390	395
Phe	Leu	Lys	Asn	Leu	Arg	Leu	Ile	Leu	Gly	Glu	Glu	Gln	Leu	Glu	Gly	405	410	415
Asn	Tyr	Ser	Phe	Tyr	Val	Leu	Asp	Asn	Gln	Asn	Leu	Gln	Gln	Leu	Trp	420	425	430

Asp Trp Asp His Arg Asn Leu Thr Ile Lys Ala Gly Lys Met Tyr Phe
 435 440 445
 Ala Phe Asn Pro Lys Leu Cys Val Ser Glu Ile Tyr Arg Met Glu Glu
 450 455 460
 Val Thr Gly Thr Lys Gly Arg Gln Ser Lys Gly Asp Ile Asn Thr Arg
 465 470 475 480
 Asn Asn Gly Glu Arg Ala Ser Cys Glu Ser Asp Val Leu His Phe Thr
 485 490 495
 Ser Thr Thr Thr Ser Lys Asn Arg Ile Ile Ile Thr Trp His Arg Tyr
 500 505 510
 Arg Pro Pro Asp Tyr Arg Asp Leu Ile Ser Phe Thr Val Tyr Tyr Lys
 515 520 525
 Glu Ala Pro Phe Lys Asn Val Thr Glu Tyr Asp Gly Gln Asp Ala Cys
 530 535 540
 Gly Ser Asn Ser Trp Asn Met Val Asp Val Asp Leu Pro Pro Asn Lys
 545 550 555 560
 Asp Val Glu Pro Gly Ile Leu Leu His Gly Leu Lys Pro Trp Thr Gln
 565 570 575
 Tyr Ala Val Tyr Val Lys Ala Val Thr Leu Thr Met Val Glu Asn Asp
 580 585 590
 His Ile Arg Gly Ala Lys Ser Glu Ile Leu Tyr Ile Arg Thr Asn Ala
 595 600 605
 Ser Val Pro Ser Ile Pro Leu Asp Val Leu Ser Ala Ser Asn Ser Ser
 610 615 620
 Ser Gln Leu Ile Val Lys Trp Asn Pro Pro Ser Leu Pro Asn Gly Asn
 625 630 635 640
 Leu Ser Tyr Tyr Ile Val Arg Trp Gln Arg Gln Pro Gln Asp Gly Tyr
 645 650 655
 Leu Tyr Arg His Asn Tyr Cys Ser Lys Asp Lys Ile Pro Ile Arg Lys
 660 665 670
 Tyr Ala Asp Gly Thr Ile Asp Ile Glu Glu Val Thr Glu Asn Pro Lys
 675 680 685
 Thr Glu Val Cys Gly Gly Glu Lys Gly Pro Cys Cys Ala Cys Pro Lys
 690 695 700
 Thr Glu Ala Glu Lys Gln Ala Glu Lys Glu Glu Ala Glu Tyr Arg Lys
 705 710 715 720
 Val Phe Glu Asn Phe Leu His Asn Ser Ile Phe Val Pro Arg Pro Glu
 725 730 735
 Arg Lys Arg Arg Asp Val Met Gln Val Ala Asn Thr Thr Met Ser Ser
 740 745 750

Arg	Ser	Arg	Asn	Thr	Thr	Ala	Ala	Asp	Thr	Tyr	Asn	Ile	Thr	Asp	Pro	755	760	765
Glu	Glu	Leu	Glu	Thr	Glu	Tyr	Pro	Phe	Phe	Glu	Ser	Arg	Val	Asp	Asn	770	775	780
Lys	Glu	Arg	Thr	Val	Ile	Ser	Asn	Leu	Arg	Pro	Phe	Thr	Leu	Tyr	Arg	785	790	795
Ile	Asp	Ile	His	Ser	Cys	Asn	His	Glu	Ala	Glu	Lys	Leu	Gly	Cys	Ser	805	810	815
Ala	Ser	Asn	Phe	Val	Phe	Ala	Arg	Thr	Met	Pro	Ala	Glu	Gly	Ala	Asp	820	825	830
Asp	Ile	Pro	Gly	Pro	Val	Thr	Trp	Glu	Pro	Arg	Pro	Glu	Asn	Ser	Ile	835	840	845
Phe	Leu	Lys	Trp	Pro	Glu	Pro	Glu	Asn	Pro	Asn	Gly	Leu	Ile	Leu	Met	850	855	860
Tyr	Glu	Ile	Lys	Tyr	Gly	Ser	Gln	Val	Glu	Asp	Gln	Arg	Glu	Cys	Val	865	870	875
Ser	Arg	Gln	Glu	Tyr	Arg	Lys	Tyr	Gly	Gly	Ala	Lys	Leu	Asn	Arg	Leu	885	890	895
Asn	Pro	Gly	Asn	Tyr	Thr	Ala	Arg	Ile	Gln	Ala	Thr	Ser	Leu	Ser	Gly	900	905	910
Asn	Gly	Ser	Trp	Thr	Asp	Pro	Val	Phe	Phe	Tyr	Val	Gln	Ala	Lys	Thr	915	920	925
Gly	Tyr	Glu	Asn	Phe	Ile	His	Leu	Ile	Ile	Ala	Leu	Pro	Val	Ala	Val	930	935	940
Leu	Leu	Ile	Val	Gly	Gly	Leu	Val	Ile	Met	Leu	Tyr	Val	Phe	His	Arg	945	950	955
Lys	Arg	Asn	Asn	Ser	Arg	Leu	Gly	Asn	Gly	Val	Leu	Tyr	Ala	Ser	Val	965	970	975
Asn	Pro	Glu	Tyr	Phe	Ser	Ala	Ala	Asp	Val	Tyr	Val	Pro	Asp	Glu	Trp	980	985	990
Glu	Val	Ala	Arg	Glu	Lys	Ile	Thr	Met	Ser	Arg	Glu	Leu	Gly	Gln	Gly	995	1000	1005
Ser	Phe	Gly	Met	Val	Tyr	Glu	Gly	Val	Ala	Lys	Gly	Val	Val	Lys	Asp	1010	1015	1020
Glu	Pro	Glu	Thr	Arg	Val	Ala	Ile	Lys	Thr	Val	Asn	Glu	Ala	Ala	Ser	1025	1030	1035
Met	Arg	Glu	Arg	Ile	Glu	Phe	Leu	Asn	Glu	Ala	Ser	Val	Met	Lys	Glu	1045	1050	1055
Phe	Asn	Cys	His	His	Val	Val	Arg	Leu	Leu	Gly	Val	Val	Ser	Gln	Gly	1060	1065	1070

Gln Pro Thr Leu Val Ile Met Glu Leu Met Thr Arg Gly Asp Leu Lys
 1075 1080 1085
 Ser Tyr Leu Arg Ser Leu Arg Pro Glu Met Glu Asn Asn Pro Val Leu
 1090 1095 1100
 Ala Pro Pro Ser Leu Ser Lys Met Ile Gln Met Ala Gly Glu Ile Ala
 1105 1110 1115 1120
 Asp Gly Met Ala Tyr Leu Asn Ala Asn Lys Phe Val His Arg Asp Leu
 1125 1130 1135
 Ala Ala Arg Asn Cys Met Val Ala Glu Asp Phe Thr Val Lys Ile Gly
 1140 1145 1150
 Asp Phe Gly Met Thr Arg Asp Ile Tyr Glu Thr Asp Tyr Tyr Arg Lys
 1155 1160 1165
 Gly Gly Lys Gly Leu Leu Pro Val Arg Trp Met Ser Pro Glu Ser Leu
 1170 1175 1180
 Lys Asp Gly Val Phe Thr Thr Tyr Ser Asp Val Trp Ser Phe Gly Val
 1185 1190 1195 1200
 Val Leu Trp Glu Ile Ala Thr Leu Ala Glu Gln Pro Tyr Gln Gly Leu
 1205 1210 1215
 Ser Asn Glu Gln Val Leu Arg Phe Val Met Glu Gly Gly Leu Leu Asp
 1220 1225 1230
 Lys Pro Asp Asn Cys Pro Asp Met Leu Phe Glu Leu Met Arg Met Cys
 1235 1240 1245
 Trp Gln Tyr Asn Pro Lys Met Arg Pro Ser Phe Leu Glu Ile Ile Ser
 1250 1255 1260
 Ser Ile Lys Glu Glu Met Glu Pro Gly Phe Arg Glu Val Ser Phe Tyr
 1265 1270 1275 1280
 Tyr Ser Glu Glu Asn Lys Leu Pro Glu Pro Glu Glu Leu Asp Leu Glu
 1285 1290 1295
 Pro Glu Asn Met Glu Ser Val Pro Leu Asp Pro Ser Ala Ser Ser Ser
 1300 1305 1310
 Ser Leu Pro Leu Pro Asp Arg His Ser Gly His Lys Ala Glu Asn Gly
 1315 1320 1325
 Pro Gly Pro Gly Val Leu Val Leu Arg Ala Ser Phe Asp Glu Arg Gln
 1330 1335 1340
 Pro Tyr Ala His Met Asn Gly Gly Arg Lys Asn Glu Arg Ala Leu Pro
 1345 1350 1355 1360
 Leu Pro Gln Ser Ser Thr Cys
 1365

<210> 11
 <211> 2394
 <212> DNA
 <213> Homo sapiens

<220>
 <223> ubiquitin-conjugating enzyme E2 variant 1
 (UBE2V1), transcript variant 2 cDNA

```

<400> 11
ttcacacggc acgacttcat cgagaccaac ggggacaact gcctggtgac actgctgccg 60
ctgctaataa tggcctacaa gttccgcacc cacagccctg aagccctgga gcagctatac 120
ccctggggagt gcttcgtctt ctgcctgata atcttcggca ccttcaccaa ccagatccac 180
aagtgggtcgc acacgtactt tgggctgcca cgctgggtca ccctcctgca ggactggcat 240
gtcatcctgc cacgtaaaaa ccatcgcata caccacgtct caccacacga gacctacttc 300
tgcataacca caggagtaaa agtccctcgc aatttccgac tgttggaaga actcgaagaa 360
ggccagaaaag gagtaggaga tggcacagtt agctgggggtc tagaagatga cgaagacatg 420
acacttacaa gatggacagg gatgataatt gggcctccaa gaacaattta tgaaaaccga 480
atatacagcc ttaaaataga atgtggacct aaatacccag aagcaccccc ctttgtaaga 540
tttgtaacaa aaattaatat gaatggagta aatagttcta atggagtggg ggacccaaga 600
gccatatcag tgctagcaaa atggcagaat tcatatagca tcaaagttgt cctgcaagag 660
cttcggcgcc taatgatgtc taaagaaaat atgaaactcc ctcagccgcc cgaaggacag 720
tggtacagca attaatcaaa aagaaaaacc acaggccctt ccccttcccc ccaattcgat 780
ttaatcagtc ttcatthttc acagtagtaa atthttctaga tacgtcttgt agacctcaaa 840
gtaccggaaa ggaagctccc attcaaagga aatttatctt aagatactgt aaatgatact 900
aattttttgt ccatttgaaa tatataagtt gtgtctataac aaatcatcct gtcaagtgtg 960
accactgtcc acgtagtgtg acttctggga tcaagaaagt ctattttaa attgattcccat 1020
cataactggg ggggcacatc taactcaact gtgaaaagac acatcacaca atcaccttgc 1080
tgctgattac acggcctggg gtctctgcct tctcccttta ccctcccgcc tcccaccctc 1140
cctgcaacaa cagccctcta gcctgggggg ctgtgttagag tagatgtgaa ggtttcagg 1200
cgcagcctgt gggactagt ctagggtgtg cgggtgtttc gcctgcaccc ctgggtcctt 1260
taagtcttaa gtgatgcccc ttccaaacca tcatcctgtc cccacgctcc tccactcccg 1320
cccttgcccg aagcatagat tgtaaccctt ccaactccct ctgagattgg cttcgggtgag 1380
gaattcaggg ctttccccat atcttctctc ccccccactt tatcgagggg tgctgctttt 1440
tctccctcct cctcaagttc ctttttgac cgtcaccacc caacacctt catgacactt 1500
ccttgctttg gccagaagcc atcaggttaag gttggaaaga gcctctgacc tcccttgttt 1560
agttttggaa ccatactcac tcaactctca ccagcctggg aaatgaatat tgggtcctca 1620
gccctgccac cctctgctgt catcagctga tgcattgttt ttagctcagg ttttgataag 1680
gtgaaaagaa tagtcaccag ggttactcag acctgccagc tctcggagtc cttggtggtt 1740
gaacttggag aaagaccgca tgaagatact tgtaagcaca catgatccct ctgaattggt 1800
ttactttcct gtaactgctt ttgcttttaa aaattgaaga agtttttaaac agggctttca 1860
tttggtcatc cttgcaatcc attggggtct agttttggaat ctgacaactg gaacaaaaag 1920
aaccttgaat ccggtgcatg ccttggtttt ggtgtgctgt ctgcttccca agatcctcag 1980
cagggattaa gaaggaaccc ggtgtgcaca gcagatcccc gaaattggtg ggcttgacct 2040
cctggcaaat tgctgctgtc ttccacttgc tgttcaggac cactaaatgc tgaaatgtgg 2100
atgcataccg aaataaaaagc aattcattgt gtactaaagg tttttttttt ttttttaatt 2160
tagtatttgt gtaaaaccac cttttgaagc agcaactatc aagtctgaaa agcaattgat 2220
gtttccatta atctttttct ggggggaaaa ccttagttct aaggatttaa catcctgtaa 2280
gtgaagttaa acataacagt attccataag cagccttttt attgtcagac cattgcctga 2340
ttttaatata ataaaaaaaa agtgtgcgtt aataaaaaaaaa aaaaaaaaaa aaaa 2394

```

<210> 12
 <211> 221
 <212> PRT
 <213> Homo sapiens

<220>
 <223> ubiquitin-conjugating enzyme E2 variant 1
 (UBE2V1), transcript variant 2

<400> 12
Met Ala Tyr Lys Phe Arg Thr His Ser Pro Glu Ala Leu Glu Gln Leu
1 5 10 15
Tyr Pro Trp Glu Cys Phe Val Phe Cys Leu Ile Ile Phe Gly Thr Phe
20 25 30
Thr Asn Gln Ile His Lys Trp Ser His Thr Tyr Phe Gly Leu Pro Arg
35 40 45
Trp Val Thr Leu Leu Gln Asp Trp His Val Ile Leu Pro Arg Lys His
50 55 60
His Arg Ile His His Val Ser Pro His Glu Thr Tyr Phe Cys Ile Thr
65 70 75 80
Thr Gly Val Lys Val Pro Arg Asn Phe Arg Leu Leu Glu Glu Leu Glu
85 90 95
Glu Gly Gln Lys Gly Val Gly Asp Gly Thr Val Ser Trp Gly Leu Glu
100 105 110
Asp Asp Glu Asp Met Thr Leu Thr Arg Trp Thr Gly Met Ile Ile Gly
115 120 125
Pro Pro Arg Thr Ile Tyr Glu Asn Arg Ile Tyr Ser Leu Lys Ile Glu
130 135 140
Cys Gly Pro Lys Tyr Pro Glu Ala Pro Pro Phe Val Arg Phe Val Thr
145 150 155 160
Lys Ile Asn Met Asn Gly Val Asn Ser Ser Asn Gly Val Val Asp Pro
165 170 175
Arg Ala Ile Ser Val Leu Ala Lys Trp Gln Asn Ser Tyr Ser Ile Lys
180 185 190
Val Val Leu Gln Glu Leu Arg Arg Leu Met Met Ser Lys Glu Asn Met
195 200 205
Lys Leu Pro Gln Pro Pro Glu Gly Gln Cys Tyr Ser Asn
210 215 220

<210> 13
<211> 1506
<212> DNA
<213> Homo sapiens

<220>
<223> NADP+ dependent aldehyde dehydrogenase 1, soluble
(ALDH1) cDNA

<400> 13
atgtcatcct caggcacgcc agacttacct gtcctactca ccgatttgaa gattcaatat 60
actaagatct tcataaacia tgaatggcat gattcagtg gtggcaagaa atttcctgtc 120
tttaatcctg caactgagga ggagctctgc caggtagaag aaggagataa ggaggatgtt 180
gacaaggcag tgaaggccgc aagacaggct tttcagattg gatctccgtg gcgtactatg 240
gatgcttccg agagggggcg actattatac aagttggctg atttaatcga aagagatcgt 300
ctgctgctgg cgacaatgga gtcaatgaat ggtggaaaac tctattccaa tgcataatctg 360
agtgatttag caggctgcat caaaacattg cgctactgtg cagggtgggc tgacaagatc 420


```

cagggccgta caataccaat tgatggaaat ttttttacat atacaagaca tgaacctatt 480
ggtgtatgtg gccaaatcat tccttggaat ttcccgttgg ttatgctcat ttggaagata 540
gggcctgcac tgagctgtgg aaacacagtg gttgtcaaac cagcagagca aactcctctc 600
actgctctcc acgtggcatc ttttaataaaa gaggcagggt ttcctcctgg agtagtgaat 660
attgttcctg gttatgggcc tacagcaggg gcagccattt cttctcacat ggatatagac 720
aaagtagcct tcacaggatc aacagagggt ggcaagttga tcaaagaagc tgccgggaaa 780
agcaatctga agagggtgac cctggagctt ggaggaaaga gcccttgcac tgtgttagct 840
gatgccgact tggacaatgc tgttgaattt gcacaccatg gggatttcta ccaccagggc 900
cagtgttgta tagccgcac caggatTTTT gtggaagaat caatttatga tgagtttgtt 960
cgaaggagtg ttgagcgggc taagaagtat atccttgga atcctctgac cccaggagtc 1020
actcaaggcc ctcagattga caaggaacaa tatgataaaa tacttgacct cattgagagt 1080
gggaagaaaag aaggggccaa actggaatgt ggaggaggcc cgtgggggaa taaaggctac 1140
tttgtccagc ccacagtgtt ctctaattgt acagatgaga tgcgcattgc caaagaggag 1200
atTTTTggac cagtgcagca aatcatgaag tttaaatctt tagatgacgt gatcaaaaga 1260
gcaaacaata ctttctatgg cttatcagca ggagtgttta ccaaagacat tgataaagcc 1320
ataacaatct cctctgctct gcaggcagga acagtgtggg tgaattgcta tggcgtggta 1380
agtgccagct gcccttttgg cggattcaag atgtctggaa atggaagaga actgggagag 1440
tacggtttcc atgaatatac agaggtcaaa acagtcacag tgaaaatctc tcagaagaac 1500
tcataa 1506

```

<210> 14

<211> 501

<212> PRT

<213> Homo sapiens

<220>

<223> NADP+ dependent aldehyde dehydrogenase 1, soluble
(ALDH1)

<400> 14

```

Met Ser Ser Ser Gly Thr Pro Asp Leu Pro Val Leu Leu Thr Asp Leu
  1              5              10              15

```

```

Lys Ile Gln Tyr Thr Lys Ile Phe Ile Asn Asn Glu Trp His Asp Ser
      20              25              30

```

```

Val Ser Gly Lys Lys Phe Pro Val Phe Asn Pro Ala Thr Glu Glu Glu
      35              40              45

```

```

Leu Cys Gln Val Glu Glu Gly Asp Lys Glu Asp Val Asp Lys Ala Val
      50              55              60

```

```

Lys Ala Ala Arg Gln Ala Phe Gln Ile Gly Ser Pro Trp Arg Thr Met
      65              70              75              80

```

```

Asp Ala Ser Glu Arg Gly Arg Leu Leu Tyr Lys Leu Ala Asp Leu Ile
      85              90              95

```

```

Glu Arg Asp Arg Leu Leu Leu Ala Thr Met Glu Ser Met Asn Gly Gly
      100             105             110

```

```

Lys Leu Tyr Ser Asn Ala Tyr Leu Ser Asp Leu Ala Gly Cys Ile Lys
      115             120             125

```

```

Thr Leu Arg Tyr Cys Ala Gly Trp Ala Asp Lys Ile Gln Gly Arg Thr
      130             135             140

```

```

Ile Pro Ile Asp Gly Asn Phe Phe Thr Tyr Thr Arg His Glu Pro Ile
      145             150             155             160

```

Gly	Val	Cys	Gly	Gln	Ile	Ile	Pro	Trp	Asn	Phe	Pro	Leu	Val	Met	Leu	165	170	175	
Ile	Trp	Lys	Ile	Gly	Pro	Ala	Leu	Ser	Cys	Gly	Asn	Thr	Val	Val	Val	180	185	190	
Lys	Pro	Ala	Glu	Gln	Thr	Pro	Leu	Thr	Ala	Leu	His	Val	Ala	Ser	Leu	195	200	205	
Ile	Lys	Glu	Ala	Gly	Phe	Pro	Pro	Gly	Val	Val	Asn	Ile	Val	Pro	Gly	210	215	220	
Tyr	Gly	Pro	Thr	Ala	Gly	Ala	Ala	Ile	Ser	Ser	His	Met	Asp	Ile	Asp	225	230	235	240
Lys	Val	Ala	Phe	Thr	Gly	Ser	Thr	Glu	Val	Gly	Lys	Leu	Ile	Lys	Glu	245	250	255	
Ala	Ala	Gly	Lys	Ser	Asn	Leu	Lys	Arg	Val	Thr	Leu	Glu	Leu	Gly	Gly	260	265	270	
Lys	Ser	Pro	Cys	Ile	Val	Leu	Ala	Asp	Ala	Asp	Leu	Asp	Asn	Ala	Val	275	280	285	
Glu	Phe	Ala	His	His	Gly	Val	Phe	Tyr	His	Gln	Gly	Gln	Cys	Cys	Ile	290	295	300	
Ala	Ala	Ser	Arg	Ile	Phe	Val	Glu	Glu	Ser	Ile	Tyr	Asp	Glu	Phe	Val	305	310	315	320
Arg	Arg	Ser	Val	Glu	Arg	Ala	Lys	Lys	Tyr	Ile	Leu	Gly	Asn	Pro	Leu	325	330	335	
Thr	Pro	Gly	Val	Thr	Gln	Gly	Pro	Gln	Ile	Asp	Lys	Glu	Gln	Tyr	Asp	340	345	350	
Lys	Ile	Leu	Asp	Leu	Ile	Glu	Ser	Gly	Lys	Lys	Glu	Gly	Ala	Lys	Leu	355	360	365	
Glu	Cys	Gly	Gly	Gly	Pro	Trp	Gly	Asn	Lys	Gly	Tyr	Phe	Val	Gln	Pro	370	375	380	
Thr	Val	Phe	Ser	Asn	Val	Thr	Asp	Glu	Met	Arg	Ile	Ala	Lys	Glu	Glu	385	390	395	400
Ile	Phe	Gly	Pro	Val	Gln	Gln	Ile	Met	Lys	Phe	Lys	Ser	Leu	Asp	Asp	405	410	415	
Val	Ile	Lys	Arg	Ala	Asn	Asn	Thr	Phe	Tyr	Gly	Leu	Ser	Ala	Gly	Val	420	425	430	
Phe	Thr	Lys	Asp	Ile	Asp	Lys	Ala	Ile	Thr	Ile	Ser	Ser	Ala	Leu	Gln	435	440	445	
Ala	Gly	Thr	Val	Trp	Val	Asn	Cys	Tyr	Gly	Val	Val	Ser	Ala	Gln	Cys	450	455	460	
Pro	Phe	Gly	Gly	Phe	Lys	Met	Ser	Gly	Asn	Gly	Arg	Glu	Leu	Gly	Glu	465	470	475	480

Tyr Gly Phe His Glu Tyr Thr Glu Val Lys Thr Val Thr Val Lys Ile
485 490 495

Ser Gln Lys Asn Ser
500

<210> 15
<211> 2282
<212> DNA
<213> Homo sapiens

<220>
<223> similar to pyruvate kinase, muscle cDNA

<400> 15
ggctgaggca gtggctcctt gcacagcagc tgcacgcgcc gtggctccgg atctcttcgt 60
ctttgcagcg tagcccgagt cggtcagcgc cggaggacct cagcagccat gtcgaagccc 120
catagtgaag ccgggactgc cttcattcag acccagcagc tgcacgcagc catggctgac 180
acattccttg agcacatgtg ccgcctggac attgattcac caccatcac agcccgaac 240
actggcatca tctgtaccat tggcccagct tcccgatcag tggagacgtt gaaggagatg 300
attaagtctg gaatgaatgt ggctcgtctg aacttctctc atggaactca tgagtaccat 360
gcgagagacca tcaagaatgt gcgcacagcc acggaaagct ttgcttctga ccccatcctc 420
taccggcccg ttgctgtggc tctagacact aaaggacctg agatccgaac tgggctcatc 480
aagggcagcg gcaactgcaga ggtggagctg aagaaggag ccactctcaa aatcacgctg 540
gataacgcct acatggaaaa gtgtgacgag aacatcctgt ggctggacta caagaacatc 600
tgcaaggttg tggaaagtgg cagcaagatc tacgtggatg atgggcttat ttctctccag 660
gtgaagcaga aaggtgccga cttcctgggtg acggaggttg aaaatggttg ctctctgggc 720
agcaagaagg gtgtgaacct tcctggggct gctgtggact tgcctgctgt gtcggagaag 780
gacatccagg atctgaagtt tggggtcgag caggatgttg atatggtgtt tgcgtcattc 840
atccgcaagg catctgatgt ccatgaagtt aggaaggtcc tgggagagaa gggaaagaac 900
atcaagatta tcagcaaaat cgagaatcat gaggggggtt ggaggtttga tgaatcctg 960
gaggccagtg atgggatcat ggtggctcgt ggtgatctag gcattgagat tcctgcagag 1020
aaggtcttcc ttgctcagaa gatgatgatt ggacggtgca accgagctgg gaagcctgtc 1080
atctgtgcta ctcagatgct ggagagcatg atcaagaagc cccgccccac tcgggctgaa 1140
ggcagtgatg tggccaatgc agtcctggat ggagccgact gcatcatgct gtctggagaa 1200
acagccaaag gggactatcc tctggagget gtgcgcagtc agcacctgat tgcccgtgag 1260
gcagaggctg ccatctacca cttgcaatta tttgaggaa cccgcccgtt ggcgccatt 1320
accagcgacc ccacagaagc caccgccgtg ggtgccgttg aggcctcctt caagtgtgtc 1380
agtggggcca taatcgtcct caccaagtct ggcaggtctg ctcaccaggt ggccagatac 1440
cgcccacgtg ccccatcat tctgtgac ccgaatcccc agacagctcg tcaggccac 1500
ctgtaccgtg gcatcttccc tgtgtgtgac aaggaccag tccaggaggc ctgggctgag 1560
gacgtggacc tccgggtgaa ctttgccatg aatgttgga aggcccgagg cttcttcaag 1620
aaggagatg tggtcattgt gctgaccgga tggcgccctg gctccggctt caccaacacc 1680
atgcgtgttg ttcctgtgcc gtgatggacc ccagagcccc tcctccagcc cctgtcccac 1740
ccccttcccc cagcccatcc attaggccag caacgcttgt agaactcact ctgggctgta 1800
acgtggcact ggtaggttgg gacaccaggg aagaagatca acgcctcact gaaacatggc 1860
tgtgtttgca gcctgtcta gtgggacagc ccagagcctg gctgcccac atgtggcccc 1920
acccaatcaa gggaagaagg aggaatgctg gactggaggc cctgggagcc agatggcaag 1980
agggtgacag cttcctttcc tgtgtgtact ctgtccagtt ccttttagaaa aaatggatgc 2040
ccagaggact cccaaccctg gcttggggtc aagaaacagc cagcaagagt taggggcctt 2100
agggcactgg gctgtgttgc cattgaagcc gactctggcc ctggccctta cttgttctc 2160
tagctctcta ggcctctcca gtttgacac gtccccaccc tcactcagc tgtcctgcag 2220
caaacactcc accctccacc ttccattttc cccactact gcagcacctc caggcctgtt 2280
gc 2282

<210> 16
<211> 531
<212> PRT
<213> Homo sapiens

<220>

<223> similar to pyruvate kinase, muscle

<400> 16

Met	Ser	Lys	Pro	His	Ser	Glu	Ala	Gly	Thr	Ala	Phe	Ile	Gln	Thr	Gln	
1				5					10					15		
Gln	Leu	His	Ala	Ala	Met	Ala	Asp	Thr	Phe	Leu	Glu	His	Met	Cys	Arg	
			20					25					30			
Leu	Asp	Ile	Asp	Ser	Pro	Pro	Ile	Thr	Ala	Arg	Asn	Thr	Gly	Ile	Ile	
		35					40					45				
Cys	Thr	Ile	Gly	Pro	Ala	Ser	Arg	Ser	Val	Glu	Thr	Leu	Lys	Glu	Met	
	50					55					60					
Ile	Lys	Ser	Gly	Met	Asn	Val	Ala	Arg	Leu	Asn	Phe	Ser	His	Gly	Thr	
65					70					75					80	
His	Glu	Tyr	His	Ala	Glu	Thr	Ile	Lys	Asn	Val	Arg	Thr	Ala	Thr	Glu	
				85					90					95		
Ser	Phe	Ala	Ser	Asp	Pro	Ile	Leu	Tyr	Arg	Pro	Val	Ala	Val	Ala	Leu	
			100					105					110			
Asp	Thr	Lys	Gly	Pro	Glu	Ile	Arg	Thr	Gly	Leu	Ile	Lys	Gly	Ser	Gly	
		115					120					125				
Thr	Ala	Glu	Val	Glu	Leu	Lys	Lys	Gly	Ala	Thr	Leu	Lys	Ile	Thr	Leu	
	130					135					140					
Asp	Asn	Ala	Tyr	Met	Glu	Lys	Cys	Asp	Glu	Asn	Ile	Leu	Trp	Leu	Asp	
145					150					155					160	
Tyr	Lys	Asn	Ile	Cys	Lys	Val	Val	Glu	Val	Gly	Ser	Lys	Ile	Tyr	Val	
			165					170					175			
Asp	Asp	Gly	Leu	Ile	Ser	Leu	Gln	Val	Lys	Gln	Lys	Gly	Ala	Asp	Phe	
			180					185					190			
Leu	Val	Thr	Glu	Val	Glu	Asn	Gly	Gly	Ser	Leu	Gly	Ser	Lys	Lys	Gly	
		195					200					205				
Val	Asn	Leu	Pro	Gly	Ala	Ala	Val	Asp	Leu	Pro	Ala	Val	Ser	Glu	Lys	
	210					215					220					
Asp	Ile	Gln	Asp	Leu	Lys	Phe	Gly	Val	Glu	Gln	Asp	Val	Asp	Met	Val	
225					230					235					240	
Phe	Ala	Ser	Phe	Ile	Arg	Lys	Ala	Ser	Asp	Val	His	Glu	Val	Arg	Lys	
				245					250					255		
Val	Leu	Gly	Glu	Lys	Gly	Lys	Asn	Ile	Lys	Ile	Ile	Ser	Lys	Ile	Glu	
			260					265					270			
Asn	His	Glu	Gly	Val	Arg	Arg	Phe	Asp	Glu	Ile	Leu	Glu	Ala	Ser	Asp	
		275					280					285				
Gly	Ile	Met	Val	Ala	Arg	Gly	Asp	Leu	Gly	Ile	Glu	Ile	Pro	Ala	Glu	
	290					295					300					

Lys Val Phe Leu Ala Gln Lys Met Met Ile Gly Arg Cys Asn Arg Ala
 305 310 315 320
 Gly Lys Pro Val Ile Cys Ala Thr Gln Met Leu Glu Ser Met Ile Lys
 325 330 335
 Lys Pro Arg Pro Thr Arg Ala Glu Gly Ser Asp Val Ala Asn Ala Val
 340 345 350
 Leu Asp Gly Ala Asp Cys Ile Met Leu Ser Gly Glu Thr Ala Lys Gly
 355 360 365
 Asp Tyr Pro Leu Glu Ala Val Arg Met Gln His Leu Ile Ala Arg Glu
 370 375 380
 Ala Glu Ala Ala Ile Tyr His Leu Gln Leu Phe Glu Glu Leu Arg Arg
 385 390 395 400
 Leu Ala Pro Ile Thr Ser Asp Pro Thr Glu Ala Thr Ala Val Gly Ala
 405 410 415
 Val Glu Ala Ser Phe Lys Cys Cys Ser Gly Ala Ile Ile Val Leu Thr
 420 425 430
 Lys Ser Gly Arg Ser Ala His Gln Val Ala Arg Tyr Arg Pro Arg Ala
 435 440 445
 Pro Ile Ile Ala Val Thr Arg Asn Pro Gln Thr Ala Arg Gln Ala His
 450 455 460
 Leu Tyr Arg Gly Ile Phe Pro Val Leu Cys Lys Asp Pro Val Gln Glu
 465 470 475 480
 Ala Trp Ala Glu Asp Val Asp Leu Arg Val Asn Phe Ala Met Asn Val
 485 490 495
 Gly Lys Ala Arg Gly Phe Phe Lys Lys Gly Asp Val Val Ile Val Leu
 500 505 510
 Thr Gly Trp Arg Pro Gly Ser Gly Phe Thr Asn Thr Met Arg Val Val
 515 520 525
 Pro Val Pro
 530

<210> 17

<211> 2631

<212> DNA

<213> Homo sapiens

<220>

<223> glucose-6-phosphate dehydrogenase (G6PD) cDNA

<400> 17

agggacagcc cagaggaggc gtggccacgc tgccggcgga agtggagccc tccgcgagcg 60
 cgcgaggccg ccggggcagg cggggaaacc ggacagtagg ggcggggccg ggccggcgat 120
 ggggatgcgg gagcactacg cggagctgca cccgtgcccc ccggaattgg ggatgcagag 180
 cagcggcagc gggatatggca ggcagccggc gggccggcct ccagcgcagg tgcccagagag 240
 gcaggggctg gcctgggatg cgcgcgcacc tgccctcgcc ccgccccgcc cgcacgaggg 300
 gtggtggccg aggccccgcc ccgcacgcct cgcctgaggc gggtcgcgtc agcccaggcg 360

```

cccccccg cccccgccga ttaaattgggc cggcgggggt cagcccccg aaacgggtcgt 420
acacttcggg gctgcgagcg cggagggcgga cgacgacgaa gcgcagacag cgtcatggca 480
gagcaggtgg ccctgagccg gacccaggtg tgcgggatcc tgcgggaaga gcttttccag 540
ggcgatgcct tccatcagtc ggatacacac atattcatca tcatgggtgc atcgggtgac 600
ctggccaaga agaagatcta cccaccatc tgggtggtgt tccgggatgg ccttctgccc 660
gaaaacacct tcatcatggg ctatgcccg tccgcctca cagtggctga catccgaaa 720
cagagtgagc ctttcttcaa ggccacccca gaggagaagc tcaagctgga ggacttcttt 780
gcccgcact cctatgtggc tggccagtac gatgatgcag cctcctacca gcgcctcaac 840
agccacatgg atgccctcca cctgggggtca caggccaacc gcctcttcta cctggccttg 900
cccccgaccg tctacgaggc cgtcaccaag aacattcacg agtcctgcat gagccagata 960
ggctggaacc gcatcatcgt ggagaagccc ttcgggaggg acctgcagag ctctgaccgg 1020
ctgtccaacc acatctcctc cctgttccgt gaggacaga tctaccgcat cgaccactac 1080
ctgggcaagg agatgggtgca gaacctcatg gtgctgagat ttgccaacag gatcttcggc 1140
cccattctgga accgggacaa catgcctgct gttatcctca ctttcaagga gccctttggc 1200
actgagggtc gcgggggcta tttcgatgaa tttgggatca tccgggacgt gatgcagaac 1260
cacctactgc agatgctgtg tctggtggcc atggagaagc ccgcctccac caactcagat 1320
gacgtccgtg atgagaaggt caaggtgttg aaatgcatct cagaggtgca ggccaacaat 1380
gtggtcctgg gccagtacgt ggggaacccc gatggagagg gcgaggccac caaagggtac 1440
ctggacgacc ccacggtgcc ccgcgggtcc accaccgcca cttttgcagc cgtcgtcctc 1500
tatgtggaga atgagaggtg ggatgggggt cccttcatcc tgcgtgcgg caaggccctg 1560
aacgagcgca aggccgaggt gaggctgcag ttccatgatg tggccggcga catcttccac 1620
cagcagtgca agcgcaacga gctggtgatc cgcgtgcagc ccaacgaggc cgtgtacacc 1680
aagatgatga ccaagaagcc gggcatgttc ttcaaccccg aggagtcgga gctggacctg 1740
acctacggca acagatacaa gaacgtgaag ctccctgacg cctacgagcg cctcatcctg 1800
gacgtcttct gcgggagcca gatgcacttc gtgcgcagcg acgagctccg tgaggcctgg 1860
cgtattttca cccactgct gcaccagatt gagctggaga agcccaagcc catcccctat 1920
atztatggca gccgaggccc cacggaggca gacgagctga tgaagagagt gggtttccag 1980
tatgagggca cctacaagtg ggtgaacccc cacaagctct gagccctggg caccacctc 2040
caccgccgcc acggccaccc tcttcccgcc cgcccgaccc cgagtcggga ggactccggg 2100
accattgacc tcagctgcac attcctggcc ccgggctctg gccaccctgg cccgccctc 2160
gctgctgcta ctaccgagc ccagctacat tctcagctg ccaagcactc gagaccatcc 2220
tggccccctc agaccctgcc tgagcccagg agctgagtc cctcctccac tactccagc 2280
ccaacagaag gaaggaggag ggcgccatt cgtctgtccc agagcttatt ggccactggg 2340
tctactcct gagtggggcc aggggtggag ggagggacaa gggggaggaa agggcgagc 2400
accacgtga gagaatctgc ctgtggcctt gcccgccagc ctacgtgcca cttgacattc 2460
cttgtcacca gcaacatctc gagccccctg gatgtcccct gtcccaccaa ctctgcactc 2520
catggccacc ccgtgccacc cgtaggcagc ctctctgcta taagaaaagc agacgcagca 2580
gctgggaccc ctcccaacct caatgcctg ccattaaatc cgcaaacagc c 2631

```

<210> 18

<211> 515

<212> PRT

<213> Homo sapiens

<220>

<223> glucose-6-phosphate dehydrogenase (G6PD)

<400> 18

```

Met Ala Glu Gln Val Ala Leu Ser Arg Thr Gln Val Cys Gly Ile Leu
  1             5             10             15

```

```

Arg Glu Glu Leu Phe Gln Gly Asp Ala Phe His Gln Ser Asp Thr His
      20             25             30

```

```

Ile Phe Ile Ile Met Gly Ala Ser Gly Asp Leu Ala Lys Lys Lys Ile
    35             40             45

```

```

Tyr Pro Thr Ile Trp Trp Leu Phe Arg Asp Gly Leu Leu Pro Glu Asn
    50             55             60

```

Thr	Phe	Ile	Met	Gly	Tyr	Ala	Arg	Ser	Arg	Leu	Thr	Val	Ala	Asp	Ile	65	70	75	80
Arg	Lys	Gln	Ser	Glu	Pro	Phe	Phe	Lys	Ala	Thr	Pro	Glu	Glu	Lys	Leu	85	90	95	
Lys	Leu	Glu	Asp	Phe	Phe	Ala	Arg	Asn	Ser	Tyr	Val	Ala	Gly	Gln	Tyr	100	105	110	
Asp	Asp	Ala	Ala	Ser	Tyr	Gln	Arg	Leu	Asn	Ser	His	Met	Asp	Ala	Leu	115	120	125	
His	Leu	Gly	Ser	Gln	Ala	Asn	Arg	Leu	Phe	Tyr	Leu	Ala	Leu	Pro	Pro	130	135	140	
Thr	Val	Tyr	Glu	Ala	Val	Thr	Lys	Asn	Ile	His	Glu	Ser	Cys	Met	Ser	145	150	155	160
Gln	Ile	Gly	Trp	Asn	Arg	Ile	Ile	Val	Glu	Lys	Pro	Phe	Gly	Arg	Asp	165	170	175	
Leu	Gln	Ser	Ser	Asp	Arg	Leu	Ser	Asn	His	Ile	Ser	Ser	Leu	Phe	Arg	180	185	190	
Glu	Asp	Gln	Ile	Tyr	Arg	Ile	Asp	His	Tyr	Leu	Gly	Lys	Glu	Met	Val	195	200	205	
Gln	Asn	Leu	Met	Val	Leu	Arg	Phe	Ala	Asn	Arg	Ile	Phe	Gly	Pro	Ile	210	215	220	
Trp	Asn	Arg	Asp	Asn	Ile	Ala	Cys	Val	Ile	Leu	Thr	Phe	Lys	Glu	Pro	225	230	235	240
Phe	Gly	Thr	Glu	Gly	Arg	Gly	Gly	Tyr	Phe	Asp	Glu	Phe	Gly	Ile	Ile	245	250	255	
Arg	Asp	Val	Met	Gln	Asn	His	Leu	Leu	Gln	Met	Leu	Cys	Leu	Val	Ala	260	265	270	
Met	Glu	Lys	Pro	Ala	Ser	Thr	Asn	Ser	Asp	Asp	Val	Arg	Asp	Glu	Lys	275	280	285	
Val	Lys	Val	Leu	Lys	Cys	Ile	Ser	Glu	Val	Gln	Ala	Asn	Asn	Val	Val	290	295	300	
Leu	Gly	Gln	Tyr	Val	Gly	Asn	Pro	Asp	Gly	Glu	Gly	Glu	Ala	Thr	Lys	305	310	315	320
Gly	Tyr	Leu	Asp	Asp	Pro	Thr	Val	Pro	Arg	Gly	Ser	Thr	Thr	Ala	Thr	325	330	335	
Phe	Ala	Ala	Val	Val	Leu	Tyr	Val	Glu	Asn	Glu	Arg	Trp	Asp	Gly	Val	340	345	350	
Pro	Phe	Ile	Leu	Arg	Cys	Gly	Lys	Ala	Leu	Asn	Glu	Arg	Lys	Ala	Glu	355	360	365	
Val	Arg	Leu	Gln	Phe	His	Asp	Val	Ala	Gly	Asp	Ile	Phe	His	Gln	Gln	370	375	380	

Cys Lys Arg Asn Glu Leu Val Ile Arg Val Gln Pro Asn Glu Ala Val
 385 390 395 400
 Tyr Thr Lys Met Met Thr Lys Lys Pro Gly Met Phe Phe Asn Pro Glu
 405 410 415
 Glu Ser Glu Leu Asp Leu Thr Tyr Gly Asn Arg Tyr Lys Asn Val Lys
 420 425 430
 Leu Pro Asp Ala Tyr Glu Arg Leu Ile Leu Asp Val Phe Cys Gly Ser
 435 440 445
 Gln Met His Phe Val Arg Ser Asp Glu Leu Arg Glu Ala Trp Arg Ile
 450 455 460
 Phe Thr Pro Leu Leu His Gln Ile Glu Leu Glu Lys Pro Lys Pro Ile
 465 470 475 480
 Pro Tyr Ile Tyr Gly Ser Arg Gly Pro Thr Glu Ala Asp Glu Leu Met
 485 490 495
 Lys Arg Val Gly Phe Gln Tyr Glu Gly Thr Tyr Lys Trp Val Asn Pro
 500 505 510
 His Lys Leu
 515

<210> 19
 <211> 1564
 <212> DNA
 <213> Homo sapiens

<220>
 <223> HCDR-3, proliferation-associated 2G4 (PA2G4)
 (protease) cDNA

<400> 19
 ctttcgctcg cccctctcctc gaggatcgag gggactctga ccacagcctg tggctgggaa 60
 gggagacaga ggcggcggcg gctcagggga aacgaggctg cagtgggtgt agtaggaaga 120
 tgtcgggcga ggacgagcaa caggagcaaa ctatcgctga ggacctgggtc gtgaccaagt 180
 ataagatggg gggcgacatc gccaacaggg tacttcgggtc cttgggtggaa gcatctagct 240
 caggtgtgtc ggtactgagc ctgtgtgaga aaggtgatgc catgattatg gaagaaacag 300
 ggaaaatctt caagaaagaa aaggaaatga agaaagggtat tgcttttccc accagcattt 360
 cggtaaataa ctgtgtatgt cacttctccc ctttgaagag cgaccaggat tatattctca 420
 aggaagggtga cttggtaaaa attgaccttg ggggtccatgt ggatggcttc atcgctaatt 480
 tagctcacac ttttgtgggt gatgtagctc aggggaccca agtaacaggg aggaaagcag 540
 atgttattaa ggcagctcac ctttgtgctg aagctgccct acgcctgggtc aaacctggaa 600
 atcagaacac acaagtga caagcctgga acaaagttgc ccactcattt aactgcacgc 660
 caatagaagg tatgtgttca caccagttga agcagcatgt catcgatgga gaaaaacca 720
 ttatccagaa tcccacagac cagcagaaga aggaccatga aaaagctgaa tttgaggtac 780
 atgaagtata tgctgtggat gttctcgtca gctcaggaga gggcaaggcc aaggatgcag 840
 gacagagaac cactatttac aaacgagacc cctctaaaca gtatggactg aaaatgaaaa 900
 cttcacgtgc cttcttcagt gaggtggaaa ggcgttttga tgccatgccg tttactttaa 960
 gagcatttga agatgagaag aaggctcgga tgggtgtggt ggagtgcgcc aaacatgaac 1020
 tgctgcaacc atttaattgt ctctatgaga aggaggggtga atttgttgcc cagtttaaat 1080
 ttacagttct gctcatgccc aatggcccca tgcggataac cagtgggtccc ttcgagcctg 1140
 acctctacaa gtctgagatg gaggtccagg atgcagagct aaaggccctc ctccagagtt 1200
 ctgcaagtgc aaaaaccag aaaaagaaaa aaaagaaggc ctccaagact gcagagaatg 1260
 ccaccagtgg ggaaacatta gaagaaaatg aagctgggga ctgaggtggg tcccattctc 1320
 ccagcttgct gctcctgcct catccccttc ccaccaaacc ccagactctg tgaagtgcag 1380

ttctttctcca cctaggaccg ccagcagagc ggggggatct ccctgcccc accccagttc 1440
 cccaacccac tcccttccaa caacaaccag ctccaactga ctctggtctt gggaggtgag 1500
 gcttcccaac cacggaagac tacttttaat gaaaaaaga aattgaataa taaaatcagg 1560
 agtc 1564

<210> 20

<211> 394

<212> PRT

<213> Homo sapiens

<220>

<223> HCDR-3, proliferation-associated 2G4 (PA2G4)
 (protease)

<400> 20

Met	Ser	Gly	Glu	Asp	Glu	Gln	Gln	Glu	Gln	Thr	Ile	Ala	Glu	Asp	Leu	1	5	10	15
Val	Val	Thr	Lys	Tyr	Lys	Met	Gly	Gly	Asp	Ile	Ala	Asn	Arg	Val	Leu	20	25	30	
Arg	Ser	Leu	Val	Glu	Ala	Ser	Ser	Ser	Gly	Val	Ser	Val	Leu	Ser	Leu	35	40	45	
Cys	Glu	Lys	Gly	Asp	Ala	Met	Ile	Met	Glu	Glu	Thr	Gly	Lys	Ile	Phe	50	55	60	
Lys	Lys	Glu	Lys	Glu	Met	Lys	Lys	Gly	Ile	Ala	Phe	Pro	Thr	Ser	Ile	65	70	75	80
Ser	Val	Asn	Asn	Cys	Val	Cys	His	Phe	Ser	Pro	Leu	Lys	Ser	Asp	Gln	85	90	95	
Asp	Tyr	Ile	Leu	Lys	Glu	Gly	Asp	Leu	Val	Lys	Ile	Asp	Leu	Gly	Val	100	105	110	
His	Val	Asp	Gly	Phe	Ile	Ala	Asn	Val	Ala	His	Thr	Phe	Val	Val	Asp	115	120	125	
Val	Ala	Gln	Gly	Thr	Gln	Val	Thr	Gly	Arg	Lys	Ala	Asp	Val	Ile	Lys	130	135	140	
Ala	Ala	His	Leu	Cys	Ala	Glu	Ala	Ala	Leu	Arg	Leu	Val	Lys	Pro	Gly	145	150	155	160
Asn	Gln	Asn	Thr	Gln	Val	Thr	Glu	Ala	Trp	Asn	Lys	Val	Ala	His	Ser	165	170	175	
Phe	Asn	Cys	Thr	Pro	Ile	Glu	Gly	Met	Leu	Ser	His	Gln	Leu	Lys	Gln	180	185	190	
His	Val	Ile	Asp	Gly	Glu	Lys	Thr	Ile	Ile	Gln	Asn	Pro	Thr	Asp	Gln	195	200	205	
Gln	Lys	Lys	Asp	His	Glu	Lys	Ala	Glu	Phe	Glu	Val	His	Glu	Val	Tyr	210	215	220	
Ala	Val	Asp	Val	Leu	Val	Ser	Ser	Gly	Glu	Gly	Lys	Ala	Lys	Asp	Ala	225	230	235	240


```

catgaaatct acatatgaac aggtggacct gattggtaaa aagactcaga aaacggcaat 1260
aactgtggag catctggcta ttaagtgcc aatggactcag agggcagcag ttattgggga 1320
tgtcatccga gtatatagt gtcataaagg acgcactatc atctttttgtg aaaccaagaa 1380
agaagcccag gagctgtccc agaattcagc tataaagcag gatgctcagt ccttgcattg 1440
agacattcca cagaagcaaa gggaaatcac cctgaaaggt tttagaaatg gtagtttttg 1500
agtttttggtg gcaaccaatg ttgctgcacg tgggttagac atccctgagg ttgatttggt 1560
tatacaaagc tctccacca aagggatgtag agtctacat tcatcgatcc gggcggacag 1620
gcagagctgg aaggacgggg gtgtgcatct gcttttatca gcacaaggaa gaatatcagt 1680
tagtacaagt ggagcaaaaa gcgggaatta agttcaaacg aataggtgtt ctttctgcaa 1740
cagaaataat aaaagcttcc agcaaagatg ccatcaggct tttggattcc gtgcctccca 1800
ctgccattag tcacttcaaa caatcagctg agaagctgat agaggagaag ggagctgttg 1860
aagctctggc agcagcactg gcccatattt cagggtgccac gtccgtagac cagcgtcctt 1920
tgatcaactc aaatgtgggt tttgtgacca tgatcttgca gtgctcaatt gaaatgcca 1980
atattagttt tgcttggaaa gaacttaaa agcagctggg cgaggagatt gattccaaag 2040
tgaagggaat ggtttttctc aaaggaaaagc tgggtgtttg ctttgatgta cctaccgcat 2100
cagtaacaga aatacaggag aaatggcatg attcacgacg ctggcagctc tctgtggcca 2160
cagagcaacc agaactggaa ggaccacggg aaggatatgg aggcttcagg ggacagcggg 2220
aaggcagtcg aggcttcagg ggacagcggg acggaaacag aagattcaga ggacagcggg 2280
aaggcagtag aggcccga ga ggacagcgtat caggaggtgg caacaaaagt aacagatccc 2340
aaaacaaagg ccagaagcgg agtttcagta aagcatttg tcaataatta gaaatagaag 2400
atttatatag caaaaagaga atgatgtttg gcaatataga actgaacatt atttttcatg 2460
caaagttaaa agcacattgt gcctcctttt gaccacttgc caagtccctg tctctttcag 2520
acacagacaa gcttcattta aattattttc tctgatcatt atcattttata actttattgt 2580
tacttcttca tcagtttttc cttttgaaag gtgtatgaat tcattacttt tttattctaa 2640
tgtattatct gtagattaga agataaaatc aagcatgtat ctgcctatac tttgtgagtt 2700
cacctgtctt tatactcaaa agtgtccctt aatagtgtcc tccctgaaa taaataccta 2760
agggagtgta acagtctctg gaggaccact ttgagccttt ggaagttaag gtttcctcag 2820
ccacctgccg aacagtttct catgtgtgtc tattatttgt ctactgagac ttaatactga 2880
gcaatgtttt gaaacaagat ttcaaactaa tctgggttgt aatacagttt ataccagtgt 2940
atgctctaga cttggaagat gtagtatgtt tgatgtggat tacctatact tatgttcgtt 3000
ttgatcatt tttagcttct cattataagg tgattcatgc tttagtgaat tcttcataga 3060
tagtatatat aaaagtacat tttaatagaa agccagggtt ttaaggaatt tcacatgtat 3120
aagggtggctc catagcttta tttgtaagta ggctggataa atgggtgctta aatggtaatg 3180
tactccactt cttcctattg gaagattaac attatttacc aagaaggact taaggagta 3240
gggggcgcag attagcattg ctcaagagta tgt 3273

```

<210> 22

<211> 558

<212> PRT

<213> Homo sapiens

<220>

<223> DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 21
(DDX21) (RNA helicase II)

<400> 22

Met Pro Gly Lys Leu Arg Ser Asp Ala Gly Leu Glu Ser Asp Thr Ala
1 5 10 15

Met Lys Lys Gly Glu Thr Leu Arg Lys Gln Thr Glu Glu Lys Glu Lys
20 25 30

Lys Glu Lys Pro Lys Ser Asp Lys Thr Glu Glu Ile Ala Glu Glu Glu
35 40 45

Glu Thr Val Phe Pro Lys Ala Lys Gln Val Lys Lys Lys Ala Glu Pro
50 55 60

Ser Glu Val Asp Met Asn Ser Pro Lys Ser Lys Lys Ala Lys Lys Lys
65 70 75 80

Glu	Glu	Pro	Ser	Gln	Asn	Asp	Ile	Ser	Pro	Lys	Thr	Lys	Ser	Leu	Arg	85	90	95
Lys	Lys	Lys	Glu	Pro	Ile	Glu	Lys	Lys	Val	Val	Ser	Ser	Lys	Thr	Lys	100	105	110
Lys	Val	Thr	Lys	Asn	Glu	Glu	Pro	Ser	Glu	Glu	Glu	Ile	Asp	Ala	Pro	115	120	125
Lys	Pro	Lys	Lys	Met	Lys	Lys	Glu	Lys	Glu	Met	Asn	Gly	Glu	Thr	Arg	130	135	140
Glu	Lys	Ser	Pro	Lys	Leu	Lys	Asn	Gly	Phe	Pro	His	Pro	Glu	Pro	Asp	145	150	155
Cys	Asn	Pro	Ser	Glu	Ala	Ala	Ser	Glu	Glu	Ser	Asn	Ser	Glu	Ile	Glu	165	170	175
Gln	Glu	Ile	Pro	Val	Glu	Gln	Lys	Glu	Gly	Ala	Phe	Ser	Asn	Phe	Pro	180	185	190
Ile	Ser	Glu	Glu	Thr	Ile	Lys	Leu	Leu	Lys	Gly	Arg	Gly	Val	Thr	Phe	195	200	205
Leu	Phe	Pro	Ile	Gln	Ala	Lys	Thr	Phe	His	His	Val	Tyr	Ser	Gly	Lys	210	215	220
Asp	Leu	Ile	Ala	Gln	Ala	Arg	Thr	Gly	Thr	Gly	Lys	Thr	Phe	Ser	Phe	225	230	235
Ala	Ile	Pro	Leu	Ile	Glu	Lys	Leu	His	Gly	Glu	Leu	Gln	Asp	Arg	Lys	245	250	255
Arg	Gly	Arg	Ala	Pro	Gln	Val	Leu	Val	Leu	Ala	Pro	Thr	Arg	Glu	Leu	260	265	270
Ala	Asn	Gln	Val	Ser	Lys	Asp	Phe	Ser	Asp	Ile	Thr	Lys	Lys	Leu	Ser	275	280	285
Val	Ala	Cys	Phe	Tyr	Gly	Gly	Thr	Pro	Tyr	Gly	Gly	Gln	Phe	Glu	Arg	290	295	300
Met	Arg	Asn	Gly	Ile	Asp	Ile	Leu	Val	Gly	Thr	Pro	Gly	Arg	Ile	Lys	305	310	315
Asp	His	Ile	Gln	Asn	Gly	Lys	Leu	Asp	Leu	Thr	Lys	Leu	Lys	His	Val	325	330	335
Val	Leu	Asp	Glu	Val	Asp	Gln	Met	Leu	Asp	Met	Gly	Phe	Ala	Asp	Gln	340	345	350
Val	Glu	Glu	Ile	Leu	Ser	Val	Ala	Tyr	Lys	Lys	Asp	Ser	Glu	Asp	Asn	355	360	365
Pro	Gln	Thr	Leu	Leu	Phe	Ser	Ala	Thr	Cys	Pro	His	Trp	Val	Phe	Asn	370	375	380
Val	Ala	Lys	Lys	Tyr	Met	Lys	Ser	Thr	Tyr	Glu	Gln	Val	Asp	Leu	Ile	385	390	395

Gly Lys Lys Thr Gln Lys Thr Ala Ile Thr Val Glu His Leu Ala Ile
 405 410 415
 Lys Cys His Trp Thr Gln Arg Ala Ala Val Ile Gly Asp Val Ile Arg
 420 425 430
 Val Tyr Ser Gly His Gln Gly Arg Thr Ile Ile Phe Cys Glu Thr Lys
 435 440 445
 Lys Glu Ala Gln Glu Leu Ser Gln Asn Ser Ala Ile Lys Gln Asp Ala
 450 455 460
 Gln Ser Leu His Gly Asp Ile Pro Gln Lys Gln Arg Glu Ile Thr Leu
 465 470 475 480
 Lys Gly Phe Arg Asn Gly Ser Phe Gly Val Leu Val Ala Thr Asn Val
 485 490 495
 Ala Ala Arg Gly Leu Asp Ile Pro Glu Val Asp Leu Val Ile Gln Ser
 500 505 510
 Ser Pro Pro Lys Gly Cys Arg Val Leu His Ser Ser Ile Arg Ala Asp
 515 520 525
 Arg Gln Ser Trp Lys Asp Gly Gly Val His Leu Leu Leu Ser Ala Gln
 530 535 540
 Gly Arg Ile Ser Val Ser Thr Ser Gly Ala Lys Ser Gly Asn
 545 550 555

<210> 23

<211> 2005

<212> DNA

<213> Homo sapiens

<220>

<223> ARK2, serine/threonine kinase 15 (STK15) cDNA

<400> 23

cttgggtcct tgggtcgcag gcatcatgga ccgatctaaa gaaaactgca tttcaggacc 60
 tggttaaggct acagctccag ttggagggtcc aaaacgtgtt ctcgtgactc agcaatttcc 120
 ttgtcagaat ccattacctg taaatagtgg ccagggtcag cgggtcttgt gtccttcaaa 180
 ttcttcccag cgcattcctt tgcaagcaca aaagcttggt tccagtcaca agccggttca 240
 gaatcagaag cagaagcaat tgcaggcaac cagtgtacct catcctgtct ccaggccact 300
 gaataacacc caaaagagca agcagcccct gccatcggca cctgaaaata atcctgagga 360
 ggaactggca tcaaaacaga aaaatgaaga atcaaaaaag aggcagtggg ctttggaaga 420
 ctttgaaatt ggtcgccctc tgggtaaagg aaagtgttgt aatgtttatt tggcaagaga 480
 aaagcaaagc aagtttattc tggctcttaa agtgttattt aaagctcagc tggagaaagc 540
 cggagtggag catcagctca gaagagaagt agaaatacag tcccaccttc ggcataccta 600
 tattcttaga ctgtatggtt atttccatga tgctaccaga gtctaccta tttctggaata 660
 tgcaccactt ggaacagttt atagagaact tcagaaactt tcaaagtttg atgagcagag 720
 aactgctact tatataacag aattggcaaa tgccctgtct tactgtcatt cgaagagagt 780
 tattcataga gacattaagc cagagaactt acttcttgga tcagctggag agcttaaaat 840
 tgcagathtt ggggtggtcag tacatgctcc atcttccagg aggaccactc tctgtggcac 900
 cctggactac ctgccccctg aaatgattga aggtcggatg catgatgaga aggtggatct 960
 ctggagcctt ggagttcttt gctatgaatt tttagtggg aagcctcctt ttgaggcaaa 1020
 cacataccaa gagacctaca aaagaatatc acgggttgaa ttcacattcc ctgactttgt 1080
 aacagaggga gccaggggacc tcatttcaag actggtgaag cataatccca gccagaggcc 1140
 aatgctcaga gaagtacttg aacacccttg gatcacagca aattcatcaa aaccatcaaa 1200
 ttgccaaaac aaagaatcag ctagcaaaca gtcttaggaa tcgtgcaggg ggagaaatcc 1260

```

ttgagccagg gctgccatat aacctgacag gaacatgcta ctgaagttaa ttttaccatt 1320
gactgctgcc ctcaatctag aacgctacac aagaaatatt tgttttactc agcaggtgtg 1380
ccttaacctc cctattcaga aagctccaca tcaataaaca tgacactctg aagtgaaggt 1440
agccacgaga attgtgctac ttatactggg tcataatctg gaggcaaggt tcgactgcag 1500
ccgccccgtc agcctgtgct aggcattggt tcttcacagg aggcaaatcc agagcctggc 1560
tgtggggaaa gtgaccactc tgccctgacc ccgatcagtt aaggagctgt gcaataacct 1620
tcctagtacc tgagtgtgtg tgtaacttat tgggttggtg aagcctggta aagctgttgg 1680
aatgagtatg tgattctttt taagtatgaa aataaagata tatgtacaga cttgtatttt 1740
ttctctggtg gcattccttt aggaatgctg tgtgtctgtc cggcaccctg gtaggcctga 1800
ttgggtttct agtcctcctt aaccacttat ctcccatatg agagtgtgaa aaataggaac 1860
acgtgctcta cctccattta gggatttgcg tgggatacag aagaggccat gtgtctcaga 1920
gctgttaagg gcttattttt ttaaaacatt ggagtcatag catgtgtgta aactttaaat 1980
atgcaataaa ataagtatct atgtc 2005

```

```

<210> 24
<211> 403
<212> PRT
<213> Homo sapiens

```

```

<220>
<223> ARK2, serine/threonine kinase 15 (STK15)

```

```

<400> 24
Met Asp Arg Ser Lys Glu Asn Cys Ile Ser Gly Pro Val Lys Ala Thr
  1             5             10             15

Ala Pro Val Gly Gly Pro Lys Arg Val Leu Val Thr Gln Gln Phe Pro
      20             25             30

Cys Gln Asn Pro Leu Pro Val Asn Ser Gly Gln Ala Gln Arg Val Leu
      35             40             45

Cys Pro Ser Asn Ser Ser Gln Arg Ile Pro Leu Gln Ala Gln Lys Leu
      50             55             60

Val Ser Ser His Lys Pro Val Gln Asn Gln Lys Gln Lys Gln Leu Gln
      65             70             75             80

Ala Thr Ser Val Pro His Pro Val Ser Arg Pro Leu Asn Asn Thr Gln
      85             90             95

Lys Ser Lys Gln Pro Leu Pro Ser Ala Pro Glu Asn Asn Pro Glu Glu
      100            105            110

Glu Leu Ala Ser Lys Gln Lys Asn Glu Glu Ser Lys Lys Arg Gln Trp
      115            120            125

Ala Leu Glu Asp Phe Glu Ile Gly Arg Pro Leu Gly Lys Gly Lys Phe
      130            135            140

Gly Asn Val Tyr Leu Ala Arg Glu Lys Gln Ser Lys Phe Ile Leu Ala
      145            150            155            160

Leu Lys Val Leu Phe Lys Ala Gln Leu Glu Lys Ala Gly Val Glu His
      165            170            175

Gln Leu Arg Arg Glu Val Glu Ile Gln Ser His Leu Arg His Pro Asn
      180            185            190

```

Ile Leu Arg Leu Tyr Gly Tyr Phe His Asp Ala Thr Arg Val Tyr Leu
 195 200 205
 Ile Leu Glu Tyr Ala Pro Leu Gly Thr Val Tyr Arg Glu Leu Gln Lys
 210 215 220
 Leu Ser Lys Phe Asp Glu Gln Arg Thr Ala Thr Tyr Ile Thr Glu Leu
 225 230 235 240
 Ala Asn Ala Leu Ser Tyr Cys His Ser Lys Arg Val Ile His Arg Asp
 245 250 255
 Ile Lys Pro Glu Asn Leu Leu Leu Gly Ser Ala Gly Glu Leu Lys Ile
 260 265 270
 Ala Asp Phe Gly Trp Ser Val His Ala Pro Ser Ser Arg Arg Thr Thr
 275 280 285
 Leu Cys Gly Thr Leu Asp Tyr Leu Pro Pro Glu Met Ile Glu Gly Arg
 290 295 300
 Met His Asp Glu Lys Val Asp Leu Trp Ser Leu Gly Val Leu Cys Tyr
 305 310 315 320
 Glu Phe Leu Val Gly Lys Pro Pro Phe Glu Ala Asn Thr Tyr Gln Glu
 325 330 335
 Thr Tyr Lys Arg Ile Ser Arg Val Glu Phe Thr Phe Pro Asp Phe Val
 340 345 350
 Thr Glu Gly Ala Arg Asp Leu Ile Ser Arg Leu Leu Lys His Asn Pro
 355 360 365
 Ser Gln Arg Pro Met Leu Arg Glu Val Leu Glu His Pro Trp Ile Thr
 370 375 380
 Ala Asn Ser Ser Lys Pro Ser Asn Cys Gln Asn Lys Glu Ser Ala Ser
 385 390 395 400

Lys Gln Ser

<210> 25

<211> 1584

<212> DNA

<213> Homo sapiens

<220>

<223> similar to transmembrane 4 superfamily member 1
(TM4 SF1) cDNA

<400> 25

gtggtgtttg ctttctccac cagaagggca cactttcatc taatttgggg tatcactgag 60
 ctgaagacaa agagaagggg gagaaaacct agcagaccac catgtgctat gggaagtgtg 120
 cacgatgcat cggacattct ctggtggggc tcgccctcct gtgcatcgcg gctaataattt 180
 tgctttactt tccaatggg gaaacaaagt atgcctccga aaaccacctc agccgcttcg 240
 tgtggttctt ttctggcatc gtaggaggtg gcctgctgat gctcctgcca gcatttgtct 300
 tcattggggt ggaacaggat gactgctgtg gctgctgtgg ccatgaaaac tgtggcaaac 360
 gatgtgcgat gctttcttct gtattggctg ctctcattgg aattgcagga tctggctact 420
 gtgtcattgt ggcagccctt ggcttagcag aaggaccact atgtcttgat tccctcggcc 480
 agtggaacta cacctttgcc agcaccgagg gccagtacct tctggatacc tccacatggt 540

```

ccgagtgcac tgaacccaag cacattgtgg aatggaatgt atctctgttt tctatcctct 600
tggctccttg tggaattgaa ttcattctgt gtcttattca agtaataaat ggagtgcctg 660
gaggcatatg tggcctttgc tgctctcacc aacagcaata tgactgctaa aagaaccaac 720
ccaggacaga gccacaatct tcctctatct cattgtaatt tatatatctt acttgatttc 780
atctgtaaaa ctttgtatta gtgtaacata ctccccacag tctactttta caaacgcctg 840
taaagactgg catcttcaca ggatgtcagt gtttaaattt agtaaacttc tttttgttt 900
gtttatttgt ttttgttttt tttttaggaa tgaggaaaca aaccaccctc tgggggtagt 960
ttacagactg agtgacagta ctgagtatat ctgagataaa ctctataatg ttttgataa 1020
aaataacatt ccaatcacta ttgtatatat gtgcatgtat tttttaaatt aaagatgtct 1080
agttgctttt tataagacca agaaggagaa aatccgacaa cctggaaaga tttttgtttt 1140
cactgcttgt atgatgtttc ccattcatac acctataaat ctctaacaag aggccctttg 1200
aactgccttg tgttctgtga gaaacaaata tttacttaga gtggaaggac tgattgagaa 1260
tgttccaatc caaatgaatg catcacaaact tacaatgctg ctcatgtgtg tgagtactat 1320
gagattcaaa tttttctaac atatggaaag ccttttgtcc tccaaagatg agtactaggg 1380
atcatgtgtt taaaaaaaaa aaaggctacg atgactgggc aagaagaaag atgggaaact 1440
gaataaagca gttgatcagc atcattggaa catggggacg agtgacggca ggaggaccac 1500
gaggaaatac cctcaaaact aacttgttta caacaaaata aagtattcac tacgaaaaaa 1560
aaaaaaaaaa aaaaaaaaaa aaaa 1584

```

<210> 26

<211> 202

<212> PRT

<213> Homo sapiens

<220>

<223> similar to transmembrane 4 superfamily member 1
(TM4 SF1)

<400> 26

```

Met Cys Tyr Gly Lys Cys Ala Arg Cys Ile Gly His Ser Leu Val Gly
  1             5             10             15

Leu Ala Leu Leu Cys Ile Ala Ala Asn Ile Leu Leu Tyr Phe Pro Asn
      20             25             30

Gly Glu Thr Lys Tyr Ala Ser Glu Asn His Leu Ser Arg Phe Val Trp
      35             40             45

Phe Phe Ser Gly Ile Val Gly Gly Gly Leu Leu Met Leu Leu Pro Ala
      50             55             60

Phe Val Phe Ile Gly Leu Glu Gln Asp Asp Cys Cys Gly Cys Cys Gly
      65             70             75             80

His Glu Asn Cys Gly Lys Arg Cys Ala Met Leu Ser Ser Val Leu Ala
      85             90             95

Ala Leu Ile Gly Ile Ala Gly Ser Gly Tyr Cys Val Ile Val Ala Ala
      100            105            110

Leu Gly Leu Ala Glu Gly Pro Leu Cys Leu Asp Ser Leu Gly Gln Trp
      115            120            125

Asn Tyr Thr Phe Ala Ser Thr Glu Gly Gln Tyr Leu Leu Asp Thr Ser
      130            135            140

Thr Trp Ser Glu Cys Thr Glu Pro Lys His Ile Val Glu Trp Asn Val
      145            150            155            160

```


Ser Leu Phe Ser Ile Leu Leu Ala Leu Gly Gly Ile Glu Phe Ile Leu
165 170 175

Cys Leu Ile Gln Val Ile Asn Gly Val Leu Gly Gly Ile Cys Gly Phe
180 185 190

Cys Cys Ser His Gln Gln Gln Tyr Asp Cys
195 200

<210> 27
<211> 1025
<212> DNA
<213> Homo sapiens

<220>
<223> excision repair cross-complementing rodent repair
deficiency, complementation group 1 (ERCC1) cDNA

<400> 27
ccaagaccag caggtgaggc ctcgcggcgc tgaaaccgtg aggcccgagc cacaggctcc 60
agatggaccc tgggaaggac aaagaggggg tgccccagcc ctcagggccg ccagcaagga 120
agaaatattgt gatacccctc gacgaggatg aggtccctcc tggagtggcc aagcccttat 180
tccgatctac acagagcctt cccactgtgg acacctcggc ccaggcggcc cctcagacct 240
acgccgaata tgccatctca cagcctctgg aaggggctgg ggccacgtgc cccacagggt 300
cagagcccct ggcaggagag acgccaacc aggcctgaa acccggggca aaatccaaca 360
gcatcattgt gagccctcgg cagaggggca atcccgtact gaagtctgtg cgcaatgtgc 420
cctgggaatt tggcgacgta attcccgaact atgtgctggg ccagagcacc tgtgccctgt 480
tcctcagcct ccgctaccac aacctgcacc cagactacat ccatgggagg ctgcagagcc 540
tggggaagaa ctctgccttg cgggtcctgc ttgtccagggt ggatgtgaaa gatccccagc 600
agggccctcaa ggagctggct aagatgtgta tcctggccga ctgcacattg atcctcgctt 660
ggagccccga ggaagctggg cggtaacctg agacctacaa ggcctatgag cagaaaccag 720
cggacctcct gatggagaag ctagagcagg acttcgtctc ccgggtgact gaatgtctga 780
ccaccgtgaa gtcagtcaac aaaacggaca gtcagacct cctgaccaca tttggatctc 840
tggaacagct catcgccgca tcaagagaag atctggcctt atgcccaggc ctggggccctc 900
agaaagcccg gaggtgttt gatgtcctgc acgagccctt cttgaaagta ccctgatgac 960
cccagctgcc aaggaaaacc ccagtgtaat aataaatcgt cctcccaggc caggctcctg 1020
ctggc 1025

<210> 28
<211> 297
<212> PRT
<213> Homo sapiens

<220>
<223> excision repair cross-complementing rodent repair
deficiency, complementation group 1 (ERCC1)

<400> 28
Met Asp Pro Gly Lys Asp Lys Glu Gly Val Pro Gln Pro Ser Gly Pro
1 5 10 15
Pro Ala Arg Lys Lys Phe Val Ile Pro Leu Asp Glu Asp Glu Val Pro
20 25 30
Pro Gly Val Ala Lys Pro Leu Phe Arg Ser Thr Gln Ser Leu Pro Thr
35 40 45
Val Asp Thr Ser Ala Gln Ala Ala Pro Gln Thr Tyr Ala Glu Tyr Ala
50 55 60

Ile Ser Gln Pro Leu Glu Gly Ala Gly Ala Thr Cys Pro Thr Gly Ser
 65 70 75 80
 Glu Pro Leu Ala Gly Glu Thr Pro Asn Gln Ala Leu Lys Pro Gly Ala
 85 90 95
 Lys Ser Asn Ser Ile Ile Val Ser Pro Arg Gln Arg Gly Asn Pro Val
 100 105 110
 Leu Lys Phe Val Arg Asn Val Pro Trp Glu Phe Gly Asp Val Ile Pro
 115 120 125
 Asp Tyr Val Leu Gly Gln Ser Thr Cys Ala Leu Phe Leu Ser Leu Arg
 130 135 140
 Tyr His Asn Leu His Pro Asp Tyr Ile His Gly Arg Leu Gln Ser Leu
 145 150 155 160
 Gly Lys Asn Phe Ala Leu Arg Val Leu Leu Val Gln Val Asp Val Lys
 165 170 175
 Asp Pro Gln Gln Ala Leu Lys Glu Leu Ala Lys Met Cys Ile Leu Ala
 180 185 190
 Asp Cys Thr Leu Ile Leu Ala Trp Ser Pro Glu Glu Ala Gly Arg Tyr
 195 200 205
 Leu Glu Thr Tyr Lys Ala Tyr Glu Gln Lys Pro Ala Asp Leu Leu Met
 210 215 220
 Glu Lys Leu Glu Gln Asp Phe Val Ser Arg Val Thr Glu Cys Leu Thr
 225 230 235 240
 Thr Val Lys Ser Val Asn Lys Thr Asp Ser Gln Thr Leu Leu Thr Thr
 245 250 255
 Phe Gly Ser Leu Glu Gln Leu Ile Ala Ala Ser Arg Glu Asp Leu Ala
 260 265 270
 Leu Cys Pro Gly Leu Gly Pro Gln Lys Ala Arg Arg Leu Phe Asp Val
 275 280 285
 Leu His Glu Pro Phe Leu Lys Val Pro
 290 295

<210> 29

<211> 181

<212> DNA

<213> Homo sapiens

<220>

<223> cell tracker assay functional hit from retroviral
screen FANCA partial cDNA

<400> 29

ccagtgtgct ggaaaggagg aagatatacct ggctggcact ctttcagttg acagagagtg 60
 acctcaggct ggggcggctc ctctccgtg tggccccgga tcagcacacc aggctgctgc 120
 ctttcgcttt ttacagtctt ctctctact tccatgaaga cgcggtttc cagcacagtg 180
 g 181

<210> 30
<211> 603
<212> DNA
<213> Homo sapiens

<220>
<223> cell tracker assay functional hit from retroviral
screen DDX9 partial cDNA

<400> 30
ccagtgtgct ggaaagcgcc acctcctctt cctgtgccaa agtagccagt tccataggcc 60
cccctaccac cwcctcgctg gaatccccc gatcctctgt agcctccact aggccctctg 120
tagtctcctc cagagttgcc tctaaagcca cctcgggaga ctctctata gcctccacca 180
acacctgcac catatcctgc ccgaaaggag ttggcgctgc caccatagcc tccgctacca 240
tagcctccac tgctatagcc accgcatagc ctccaccact gtaactagaa cctccccctc 300
tatatccgct tccattgtcg tatcggggcca tcttgggagg acgtggacca tctccatgcc 360
gtgtactgcc aatcataagg ttgataccag cagctgaggg tctagagatc tgacggatca 420
tgttcagcat acgttcattt acgggggtcca actggctgat gatagcaggt tgtttggtta 480
cttcaacaac caaagcctcc atggctgccc ggagaccagt gatacaggca gcagcttcat 540
gagatatttg cagtttaatc cagtcatcta caagcacaat ctgcccactt tccagcacag 600
tgg 603

<210> 31
<211> 145
<212> DNA
<213> Homo sapiens

<220>
<223> cell tracker assay functional hit from retroviral
screen IGF1R partial cDNA

<400> 31
ccagtgtggt ggaaagggag agaactgtca tttctaacct tcggcctttc acattgtacc 60
gcatcgatat ccacagctgc aaccacgagg ctgagaagct gggctgcagc gcctccaact 120
tcgtctttgc tttccagcac agtgg 145

<210> 32
<211> 269
<212> DNA
<213> Homo sapiens

<220>
<223> cell tracker assay functional hit from retroviral
screen UBEV2V1 partial cDNA

<400> 32
ccagtgtgct ggaaagggtgc ttctgggtat ttaggtccac attctatttt aaggctgtat 60
attcgggtttt cataaattgt tcttggaggg ccaattatca tccctgtcca tcttgtaaga 120
tgtcatgtct tcgtcatctt ctagacccca gctaactgtg ccatctccta ctctttctg 180
gccttcttcg agattcctcc aacagtcgga aattgcgagg gactttatac atcccagacc 240
cgtggtggct gccctttcca gcacactgg 269

<210> 33
<211> 499
<212> DNA
<213> Homo sapiens

<220>
 <223> cell tracker assay functional hit from retroviral
 screen aldehyde dehydrogenase partial cDNA

<400> 33
 ccagtgtgct ggaaaggagc aaactcctct cactgctctc cacgtggcat ctttaataaa 60
 agaggcaggg tttcctcctg gagtagtgaa tattgttcct gggtatgggc ctacagcagg 120
 ggcagccatt tcttctcaca tggatataga caaagtagcc ttcacaggat caacagaggt 180
 tggcaagttg atcaaagaag ctgccgggaa aagcaatctg aagaggggtga ccctggagct 240
 tggaggaaag agcccttgca ttgtgttagc tgatgccgac ttggacaatg ctgttggaatt 300
 tgcacaccat ggggtattct accaccaggg ccagtgttgt atagccgcat ccaggatttt 360
 tgtggaagaa tcaatttatg atgagttttg ttcgaaggag tgttgagcgg gctaagaacg 420
 tatatccttg gaaacatcct ctgacccag gagtcactca aaggccctca gattgacaag 480
 gactttccag acacagtgg 499

<210> 34
 <211> 425
 <212> DNA
 <213> Homo sapiens

<220>
 <223> cell tracker assay functional hit from retroviral
 screen pyruvate dehydrogenase partial cDNA

<400> 34
 ccagtgtgct ggaaaggctg ccacttcca ccaccttgca gatgttcttg tagtccagcc 60
 acaggatggt ctcgtcacac ttttccatgt aggcgttata cagcgtgatt ttgagagtgg 120
 ctcccttctt cagctccacc tctgcagtgc cgctgccctt gatgagccca gttcggatct 180
 caggtccttt agtgtctaga gccacagcaa cgggcccggta gaggatgggg tcagaagcaa 240
 agctttccgt ggctgtgctg acattcttga tgggtctcgc atggtactca tgagttccat 300
 gagagaagtt cagacgagcc acattcattc cagacttaat catctccttc aacgtctcca 360
 ctggatcggg aagctgggac aatggtacag atgatgccag tggtccgggc tttccagcac 420
 agtgg 425

<210> 35
 <211> 238
 <212> DNA
 <213> Homo sapiens

<220>
 <223> cell tracker assay functional hit from retroviral
 screen G6PD partial cDNA

<220>
 <221> modified_base
 <222> (1)..(238)
 <223> n = g, a, c or t

<400> 35
 ccagtgtgct ggaaactttc cagttctcca tggccaccan acacagcatc tgcagtaggt 60
 ggttctgcat cagctcccgg atgatcccaa attcatcgaa atagcccccg cgaccctcag 120
 tgccaaaggg ctccctgaag gtgaggataa cgcaggcgat gttgtcccgg ttccanattg 180
 ggccgaagat cctgttggca aatctcagca ccatgaggtt ctctttccag cacagtgg 238

<210> 36
 <211> 491
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:G3-2D8 sequence
 with C-terminus of GFP, partial BAP-1 sequence

<220>
 <221> modified_base
 <222> (69)
 <223> n = g, a, c or t

<400> 36
 gagttcgtga cgcgcgcgcg gatcactctc ggcatggacg agctgtacaa ggaggaggcc 60
 gccaaggcng gtggcagcgg tggctccagt gtgctggaaa gctaagggca gagttggtgt 120
 tctgcacgtc atcctctcgt catcctcata gtcacacctca tcatctgagt actgctgggg 180
 tgggcgggact ggaactcggc tgcggccac acctgccgcc aggtcttctt cctcctgcat 240
 ggggggacttg gcataattgt gattgtctag aaaggccggc agccgctgga caatgggagt 300
 ggggttgggg tgaaccccat tgaggctgct gcctggaggc taccactagc ttgggtttgt 360
 tgggagggct gtgggatggg gcttgtgcgc atgaaccagc cgcctcctcg caccatctgt 420
 gtggttgccc tcagaggctg caggggcctt gtttgcttcc agcaccagcg gggacctttc 480
 cagcacagtg g 491

<210> 37
 <211> 34
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:G3-2D8 sequence
 with C-terminus of GFP, partial BAP-1 sequence

<400> 37
 Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr
 1 5 10 15
 Lys Glu Glu Ala Ala Lys Ala Gly Gly Ser Gly Gly Ser Ser Val Leu
 20 25 30
 Glu Ser

<210> 38
 <211> 16
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:BstXI linker

<400> 38
 ccagtgtgct ggaaaag 16

<210> 39
 <211> 16
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:BstXI linker

<400> 39
ctttccagca cagtgg

16

<210> 40
<211> 136
<212> PRT
<213> Homo sapiens

<220>
<223> G1-2635 sequence, partial NP95 sequence

<400> 40
Ala Glu Gln Ser Cys Asp Gln Lys Leu Thr Asn Thr Asn Arg Ala Leu
1 5 10 15
Ala Leu Asn Cys Phe Ala Pro Ile Asn Asp Gln Glu Gly Ala Glu Ala
20 25 30
Lys Asp Trp Arg Ser Gly Lys Pro Val Arg Val Val Arg Asn Val Lys
35 40 45
Gly Gly Lys Asn Ser Lys Tyr Ala Pro Ala Glu Gly Asn Arg Tyr Asp
50 55 60
Gly Ile Tyr Lys Val Val Lys Tyr Trp Pro Glu Lys Gly Lys Ser Gly
65 70 75 80
Phe Leu Val Trp Arg Tyr Leu Leu Arg Arg Asp Asp Asp Glu Pro Gly
85 90 95
Pro Trp Thr Lys Glu Gly Lys Asp Arg Ile Lys Lys Leu Gly Leu Thr
100 105 110
Met Gln Tyr Pro Glu Gly Tyr Leu Glu Ala Leu Ala Asn Arg Glu Arg
115 120 125
Glu Lys Glu Asn Ser Lys Arg Glu
130 135

<210> 41
<211> 49
<212> PRT
<213> Homo sapiens

<220>
<223> G2-2F3 sequence, partial FANCA sequence

<400> 41
Arg Lys Ile Ser Trp Leu Ala Leu Phe Gln Leu Thr Glu Ser Asp Leu
1 5 10 15
Arg Leu Gly Arg Leu Leu Leu Arg Val Ala Pro Asp Gln His Thr Arg
20 25 30
Leu Leu Pro Phe Ala Phe Tyr Ser Leu Leu Ser Tyr Phe His Glu Asp
35 40 45
Ala

<210> 42
 <211> 552
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:G3-2H6 sequence
 with C-terminus of GFP, partial DDX9 sequence

<400> 42
 gagttcgtga cgcgcgcgcg gatcactctc ggcattggacg agctgtacaa ggaggaggcc 60
 gccaaggccg gtggcagcgg tggctccagt gtgctggaaa gcgccacctc ctcttccttg 120
 tccaaagtag ccagttccat aggccccct accacctcct cgctggaatc cccagatcc 180
 tctgtagcct ccactaggcc ctctgtagtc tcctccagag ttgcctctaa agccacctcg 240
 ggagactcct ctatagcctc caccaacacc tgcaccatat cctgcccgaaggaggattggc 300
 gctgccacca tagcctccgc taccatagcc tccactgcta tagccaccgc atagcctcca 360
 ccactgtaac tagaacctcc ctttctatat ccgcttccat tgctgtatcg ggccatcttg 420
 ggaggacgtg gaccatctcc atgccgtgta ctgccaatca taaggttgat accagcagct 480
 gaggggtctac agatctgacg gatcatgttc agcatatcgtt catttacggg gtccaactgg 540
 ctgatgatag ca 552

<210> 43
 <211> 117
 <212> PRT
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:G3-2H6 sequence
 with C-terminus of GFP, partial DDX9 sequence

<400> 43
 Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr
 1 5 10 15
 Lys Glu Glu Ala Ala Lys Ala Gly Gly Ser Gly Gly Ser Ser Val Leu
 20 25 30
 Glu Ser Ala Thr Ser Ser Ser Leu Ser Lys Val Ala Ser Ser Ile Gly
 35 40 45
 Pro Pro Thr Thr Ser Ser Leu Glu Ser Pro Arg Ser Ser Val Ala Ser
 50 55 60
 Thr Arg Pro Ser Val Val Ser Ser Arg Val Ala Ser Lys Ala Thr Ser
 65 70 75 80
 Gly Asp Ser Ser Ile Ala Ser Thr Asn Thr Cys Thr Ile Ser Cys Pro
 85 90 95
 Lys Gly Val Gly Ala Ala Thr Ile Ala Ser Ala Thr Ile Ala Ser Thr
 100 105 110
 Ala Ile Ala Thr Ala
 115

<210> 44
 <211> 38
 <212> PRT
 <213> Homo sapiens

<220>
 <223> G3-2H2_1 sequence, partial IGF1R sequence

<400> 44
 Glu Arg Thr Val Ile Ser Asn Leu Arg Pro Phe Thr Leu Tyr Arg Ile
 1 5 10 15
 Asp Ile His Ser Cys Asn His Glu Ala Glu Lys Leu Gly Cys Ser Ala
 20 25 30
 Ser Asn Phe Val Phe Ala
 35

<210> 45
 <211> 347
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:G3-2G2/2H2
 sequence with C-terminus of GFP, partial UBE2V1
 sequence

<400> 45
 gagttcgtga ccgccgccgg gatcactctc ggcattggacg agctgtacaa ggaggaggcc 60
 gccaaggccg gtggcagcgg tggctccagt gtgctggaaa ggtgcttctg ggtatttagg 120
 tccacattct attttaaggc tgtatattcg gttttcataa attgttcttg gaggcccaat 180
 tatcatccct gtccatcttg taagtgtcat gtcttcgtca tcttctagac ccagagtaac 240
 tgtgccatct cctactcctt tctggccttc ttcgagttct tccaacagtc ggaaattgcg 300
 agggactttt actcccgagc ccgtggtggc tctttccagc acagtgg 347

<210> 46
 <211> 75
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:G3-2G2/2H2
 sequence with C-terminus of GFP, partial UBE2V1
 sequence

<400> 46
 Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr
 1 5 10 15
 Lys Glu Glu Ala Ala Lys Ala Gly Gly Ser Gly Gly Ser Ser Val Leu
 20 25 30
 Glu Arg Cys Phe Trp Val Phe Arg Ser Thr Phe Tyr Phe Lys Ala Val
 35 40 45
 Tyr Ser Val Phe Ile Asn Cys Ser Trp Arg Pro Asn Tyr His Pro Cys
 50 55 60
 Pro Ser Cys Lys Cys His Val Phe Val Ile Phe
 65 70 75

<210> 47
<211> 177
<212> DNA
<213> Homo sapiens

<220>
<223> portion of G3_2H2 sequence

<400> 47
ggcagccacc acgggctcgg gagtaaaagt ccctcgcaat ttccgactgt tggaagaact 60
cgaagaaggc cagaaaggag taggagatgg cacagtttag tggggcttag aagatgacga 120
agacatgaca cttacaagat ggacagggat gataattggg cctccaagaa caattta 177

<210> 48
<211> 198
<212> DNA
<213> Homo sapiens

<220>
<223> portion of Ubce2v1 alternatively spliced
transcript

<400> 48
ggcagcaggg cgacgcaaga tggcagccac cacgggctcg ggagtaaaag tccctcgcaa 60
tttccgactg ttggaagaac tcgaagaagg ccagaaagga gtaggagatg gcacagtttag 120
ctgggggtcta gaagatgacg aagacatgac acttacaaga tggacaggga tgataattgg 180
gcctccaaga acaattta 198

<210> 49
<211> 112
<212> DNA
<213> Homo sapiens

<220>
<223> portion of Ubce2v1_2 alternatively spliced
transcript

<400> 49
gtcatcctgc cacgtaaaca ccatcgcac caccacgtct caccacacga gacctacttc 60
tgcacaccca caggagtaaa agtccctcgc aatttccgac tggttgaaga ac 112

<210> 50
<211> 113
<212> DNA
<213> Homo sapiens

<220>
<223> portion of Ubce2v1_1 alternatively spliced
transcript

<400> 50
aagcgtctta cctgaagtca caaagcaaac tgagtgatga aggaagactt gaacctagaa 60
aatttcactg caaaggagat aaaagtcct cgcaatttcg actggttgaa gaa 113

<210> 51
<211> 192
<212> DNA
<213> Homo sapiens

<220>
<223> portion of UbcE2v1Bs alternatively spliced
transcript

<400> 51
atggcctaca agttccgcac agtggttacag ccctggagta aaagtccctc gcaatttccg 60
actgttgga gaactcgaag aaggccagaa aggagtagga gatggcacag ttagctgggg 120
tctagaagat gacgaagaca tgacacttac aagatggaca gggatgataa ttgggcctcc 180
aagaacaatt ta 192

<210> 52
<211> 194
<212> DNA
<213> Homo sapiens

<220>
<223> portion of UbcE2v2 alternatively spliced
transcript

<400> 52
cgcgtcgggc tgcaggagaa gatggcggtc tccacaggag ttaaagttcc tcgtaatttt 60
cgcttggttg aagaacttga agaaggacaa aaaggagtag gcgacggtac agttagctgg 120
ggccttgaag atgatgaaga tatgacactt acaaggtgga caggcatgat tattgggcca 180
ccaaggacaa atta 194

<210> 53
<211> 113
<212> DNA
<213> Homo sapiens

<220>
<223> portion of UbcE2v1_2 and UbcE2v1_1 alternatively spliced
transcripts

<400> 53
gaaggccaga aaggagtagg agatggcaca gttagctggg gtctagaaga tgacgaagac 60
atgacactta caagatggac agggatgata attgggcctc caagaacaat tta 113

<210> 54
<211> 49
<212> DNA
<213> Homo sapiens

<220>
<223> portion of G3_2H2 sequence

<400> 54
cgaatatata gccttaaaaat agaatgtgga cctaaatacc cagaagcac 49

<210> 55
<211> 60
<212> DNA
<213> Homo sapiens

<220>
<223> portion of Ubce2v1, Ubce2v1_2, Ubce2v1_1 and Ubce2v1Bs
alternatively spliced transcripts

<400> 55
cgaatatata gccttaaaat agaatgtgga cctaaatacc cagaagcacc cccctttgta 60

<210> 56
<211> 60
<212> DNA
<213> Homo sapiens

<220>
<223> portion of Ubce2v2 alternatively spliced
transcript

<400> 56
agaatatata gcctgaaagt agaatgtgga cctaaatacc cagaagctcc tccgtcagtt 60

<210> 57
<211> 225
<212> PRT
<213> Homo sapiens

<220>
<223> Uch-13 sequence

<400> 57
Met Glu Gly Gln Arg Trp Leu Pro Leu Glu Ala Asn Pro Glu Val Thr
1 5 10 15
Asn Gln Phe Leu Lys Gln Leu Gly Leu His Pro Asn Trp Gln Phe Val
20 25 30
Asp Val Tyr Gly Met Asp Pro Glu Leu Leu Ser Met Val Pro Arg Pro
35 40 45
Val Cys Ala Val Leu Leu Leu Phe Pro Ile Thr Glu Lys Tyr Glu Val
50 55 60
Phe Arg Thr Glu Glu Glu Glu Lys Ile Lys Ser Gln Gly Gln Asp Val
65 70 75 80
Thr Ser Ser Val Tyr Phe Met Lys Gln Thr Ile Ser Asn Ala Cys Gly
85 90 95
Thr Ile Gly Leu Ile His Ala Ile Ala Asn Asn Lys Asp Lys Met His
100 105 110
Phe Glu Ser Gly Ser Thr Leu Lys Lys Phe Leu Glu Glu Ser Val Ser
115 120 125
Met Ser Pro Glu Glu Arg Ala Arg Tyr Leu Glu Asn Tyr Asp Ala Ile
130 135 140

Arg Val Thr His Glu Thr Ser Ala His Glu Gly Gln Thr Glu Ala Pro
 145 150 155 160
 Ser Ile Asp Glu Lys Val Asp Leu His Phe Ile Ala Leu Val His Val
 165 170 175
 Asp Gly His Leu Tyr Glu Leu Asp Gly Arg Lys Pro Phe Pro Ile Asn
 180 185 190
 His Gly Glu Thr Ser Asp Glu Thr Leu Leu Glu Asp Ala Ile Glu Val
 195 200 205
 Cys Lys Lys Phe Met Glu Arg Asp Pro Asp Glu Leu Arg Phe Asn Ala
 210 215 220
 Ile
 225

<210> 58
 <211> 231
 <212> PRT
 <213> Homo sapiens

<220>
 <223> portion of BAP-1

<400> 58
 Met Asn Lys Gly Trp Leu Glu Leu Glu Ser Asp Pro Gly Leu Phe Thr
 1 5 10 15
 Leu Leu Val Glu Asp Phe Gly Val Lys Gly Val Gln Val Glu Glu Ile
 20 25 30
 Tyr Asp Leu Gln Ser Lys Cys Gln Gly Pro Val Tyr Gly Phe Ile Phe
 35 40 45
 Leu Phe Lys Trp Ile Glu Glu Arg Arg Ser Arg Arg Lys Val Ser Thr
 50 55 60
 Leu Val Asp Asp Thr Ser Val Ile Asp Asp Asp Ile Val Asn Asn Met
 65 70 75 80
 Phe Phe Ala His Gln Leu Ile Pro Asn Ser Cys Ala Thr His Ala Leu
 85 90 95
 Leu Ser Val Leu Leu Asn Cys Ser Ser Val Asp Leu Gly Pro Thr Leu
 100 105 110
 Ser Arg Met Lys Asp Phe Thr Lys Gly Phe Ser Pro Glu Ser Lys Gly
 115 120 125
 Tyr Ala Ile Gly Asn Ala Pro Glu Leu Ala Lys Ala His Asn Ser His
 130 135 140
 Ala Arg Pro Glu Pro Arg His Leu Pro Glu Lys Gln Asn Gly Leu Ser
 145 150 155 160
 Ala Val Arg Thr Met Glu Ala Phe His Phe Val Ser Tyr Val Pro Ile
 165 170 175

Thr Gly Arg Leu Phe Glu Leu Asp Gly Leu Lys Val Tyr Pro Ile Asp
 180 185 190
 His Gly Pro Trp Gly Glu Asp Glu Glu Trp Thr Asp Lys Ala Arg Arg
 195 200 205
 Val Ile Met Glu Arg Ile Gly Leu Ala Thr Ala Gly Glu Pro Tyr His
 210 215 220
 Asp Ile Arg Phe Asn Leu Met
 225 230

<210> 59
 <211> 200
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:poly Gly
 flexible linker

<220>
 <221> MOD_RES
 <222> (6)..(200)
 <223> Gly residues from position 6 to 200 may be present
 or absent

<400> 59
 Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
 1 5 10 15
 Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
 20 25 30
 Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
 35 40 45
 Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
 50 55 60
 Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
 65 70 75 80
 Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
 85 90 95
 Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
 100 105 110
 Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
 115 120 125
 Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
 130 135 140
 Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
 145 150 155 160
 Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
 165 170 175

Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
180 185 190

Gly Gly Gly Gly Gly Gly Gly Gly
195 200

<210> 60

<211> 4

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:DEAD box
conserved motif

<400> 60

Asp Glu Ala Asp

1